ЭНЦИКЛОПЕДИЯ ЭЛЕМЕНТАРНОЙ МАТЕМАТИКИ

III

ЭНЦИКЛОПЕДИЯ ЭЛЕМЕНТАРНОЙ МАТЕМАТИКИ

III

ФУНКЦИИ И ПРЕДЕЛЫ

АКАДЕМИЯ ПЕДАГОГИЧЕСКИХ НАУК РСФСР

ЭНЦИКЛОПЕДИЯ ЭЛЕМЕНТАРНОЙ МАТЕМАТИКИ

ПОД РЕДАКЦИЕЙ
П. С. АЛЕКСАНДРОВА,
А. И. МАРКУШЕВИЧА
и А. Я. ХИНЧИНА

КНИГА ТРЕТЬЯ

ФУНКЦИИ И ПРЕДЕЛЫ
(ОСНОВЫ АНАЛИЗА)

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ТЕХНИКО-ТЕОРЕТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1952 ЛЕНИНГРАД

11-5-2

Редактор *А. З. Рывкин.* Техн. редактор *Н. Я. Мурашова.*

Корректор *А. С. Каган.*

Подписано к печати 4/IV 1952 г. Бумага 60×92¹/₁₆. 17,5 бум. л. 35 печ. л. 36,95 уч.-изд. л. 42 228 тип. знаков в печ. л. Т-02117. Тираж 50 000 экз. Цена книги 11 р. 10 к. Переплёт 2 р. Заказ № 38.

Номинал — по прейскуранту 1952 г.

2-я типография «Печатный Двор» им. А. М. Горького Главполиграфиздата при Совете Министров СССР. Ленинград, Гатчинская, 26.

ОГЛАВЛЕНИЕ

Предисловие 7

ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ ДЕЙСТВИТЕЛЬНОГО ПЕРЕМЕННОГО. ПРЕДЕЛЫ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И ФУНКЦИЙ. ОБЩЕЕ ПОНЯТИЕ ФУНКЦИИ

(*В. Л. Гончаров*)

Глава I. Общие сведения об элементарных функциях и графиках уравнений 11

§ 1. Элементарные функции 11
§ 2. Графические представления. Приёмы точечных построений . . 17
§ 3. Простейшие преобразования графиков 25
§ 4. Прямая и обратная функции 32
§ 5. Элементарное исследование функций (постановка вопроса и некоторые общие приёмы) 34

Глава II. Обзор элементарных функций и их графиков 41

§ 6. Классификация рациональных функций 41
§ 7. Целые положительные степени 42
§ 8. Многочлены первой степени (линейные функции) 45
§ 9. Многочлены (трёхчлены) второй степени 46
§ 10. Многочлены третьей степени 48
§ 11. Биквадратные многочлены 51
§ 12. Многочлены высших степеней 52
§ 13. Целые отрицательные степени 54
§ 14. Дробные линейные функции 56
§ 15. Дробные функции второй степени 58
§ 16. Дробные рациональные функции (общий случай) 64
§ 17. Алгебраические иррациональные функции 66
§ 18. Примеры исследования алгебраических функций 68
§ 19. Элементарные трансцендентные функции 78
§ 20. Показательная функция 78
§ 21. Функции, связанные с показательной 84
§ 22. Логарифмическая функция 88
§ 23. Функции, связанные с логарифмической 90
§ 24. Произвольная степенная функция 93
§ 25. Основные (целые) тригонометрические функции: синус и косинус 95
§ 26. Простые гармонические колебания 101
§ 27. Тригонометрические многочлены 105
§ 28. Многочлены Чебышева 107
§ 29. Тангенс и другие дробные тригонометрические функции . . . 111

§ 30. Представление функций, рационально зависящих от тригонометрических, через одну или две из них 116
§ 31. Примеры исследования функций, рационально зависящих от тригонометрических. Тригонометрические уравнения 121
§ 32. Обратные тригонометрические функции 128
§ 33. Исследование многочленов Чебышева. Их минимальное свойство 134

Глава III. **Пределы числовых последовательностей и пределы функций** 140

§ 34. Конечные и бесконечные числовые последовательности . . . 140
§ 35. Общее определение бесконечной числовой последовательности 149
§ 36. Теорема Больцано-Вейерштрасса о существовании предельной точки 153
§ 37. Примеры. Предел как единственная предельная точка 159
§ 38. Предел последовательности: классическое определение и основные свойства 165
§ 39. Обобщение понятия предела (пределы в «несобственном смысле») 173
§ 40. Предел функции на бесконечности 176
§ 41. Односторонний предел функции в конечной точке 180
§ 42. Двусторонний предел. Понятие непрерывности 187
§ 43. Примеры непрерывных функций 190
§ 44. Пределы при монотонном изменении. Число e 195

Глава IV. **Пределы последовательностей функций. Свойства непрерывных функций** 202

§ 45. Простая сходимость 202
§ 46. Общее понятие функции одной действительной переменной 210
§ 47. Свойства непрерывных функций 215
§ 48. Равномерная сходимость последовательности непрерывных функций 222
§ 49. Теорема Вейерштрасса-Бернштейна о приближении непрерывной функции с помощью рациональных многочленов . . . 227
§ 50. Доказательство теоремы 232
§ 51. Определение показательной функции. Продолжение непрерывной функции за пределы всюду плотного множества . . . 237
§ 52. Теорема Больцано и проблема существования однозначной обратной функции 244
§ 53. Функциональные уравнения и элементарные функции 247

Глава V. **Общее понятие функции** 254

§ 54. Соответствие между множествами 254
§ 55. Геометрические образы в многомерных пространствах 256
§ 56. Пространственные отображения 260
§ 57. Метрические пространства 264
§ 58. Понятие предела в метрическом пространстве 268
§ 59. Топологические пространства 272
§ 60. Алгебра множеств. Производное множество. Замкнутость и связность 274
§ 61. Непрерывные отображения и их свойства 279
§ 62. Гомеоморфные отображения 282
§ 63. Верхняя и нижняя границы числовых множеств или последовательностей. Верхний и нижний пределы числовых множеств или последовательностей 287

ПРОИЗВОДНЫЕ, ИНТЕГРАЛЫ И РЯДЫ

(И. П. Натансон)

Введение 299

Глава I. Производные 303

§ 1. Производная и дифференциал 303
1. Задачи, приводящие к понятию производной 303
2. Определение производной 307
3. Дифференцируемость и непрерывность. Односторонние производные 309
4. Производные простейших элементарных функций 312
5. Дифференцирование обратных функций 318
6. Правила комбинирования формул дифференцирования 320
7. Дифференциал 327
8. Производные и дифференциалы высшего порядка 333
9. Частные производные и полный дифференциал 337

§ 2. Важнейшие теоремы о производных 339
10. Теоремы Ферма и Ролля 339
11. Формулы Лагранжа и Коши. Правило Лопиталя 342
12. Формула Тейлора 346
13. Исследования П. Л. Чебышева и С. Н. Бернштейна 353

§ 3. Применение дифференциального исчисления к исследованию функций 354
14. Признаки постоянства и монотонности функции 354
15. Экстремум функции 359
16. Нахождение наибольшего и наименьшего значений функции на замкнутом промежутке 363

Глава II. Интегралы 366

§ 4. Неопределённые интегралы 366
17. Основные понятия 366
18. Интегрирование с помощью подстановки 369
19. Интегрирование по частям 371
20. Общие замечания по поводу интегрирования элементарных функций 373

§ 5. Определённые интегралы 377
21. Задачи, приводящие к понятию определённого интеграла . . 377
22. Определённый интеграл 380
23. Основные свойства интеграла 385
24. Интеграл, как функция верхнего предела 391
25. Вычисление определённого интеграла с помощью неопределённого 393
26. Формула Валлиса 398
27. Приближённое вычисление определённых интегралов 400

§ 6. Приложения интегрального исчисления 408
28. Вычисление площадей 408
29. Вычисление объёмов 411
30. Длина дуги кривой 417
31. Площадь поверхности вращения 418
32. Общие указания по поводу приложений интегрального исчисления и его связей с дифференциальным исчислением 420

Глава III. Ряды 425

§ 7. Ряды с постоянными членами 425
33. Основные понятия 425
34. Простейшие свойства рядов 429
35. Положительные ряды 431
36. Знакочередующиеся ряды 437
37. Абсолютная сходимость 440
38. Вопрос о перестановке членов ряда. Умножение рядов . . . 441

§ 8. Степенные ряды 447
39. Промежуток сходимости 447
40. Свойства суммы степенного ряда 452
41. Разложение логарифма и составление таблиц логарифмов . . 457
42. Разложение арктангенса и вычисление π 465
43. Общие замечания по поводу разложения функций в степенные ряды 469
44. Биномиальный ряд 472
45. Очерк аналитической теории тригонометрических функций . 481

ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО

(*В. Л. Гончаров*)

§ 1. Рациональные функции 493
§ 2. Пределы. Ряды 496
§ 3. Показательная функция. Синус и косинус 500
§ 4. Выражение тригонометрических функций через показательную 504
§ 5. Гиперболические и тригонометрические функции 507
§ 6. Логарифм 508
§ 7. Произвольная степень 510
§ 8. Обратные тригонометрические и гиперболические функции . 511
§ 9. Производная 513
§ 10. Интеграл 517
§ 11. Приближение функций многочленами 523
§ 12. Первообразная функция 526
§ 13. Интеграл Коши 532
§ 14. Понятие аналитической функции 536
§ 15. Свойства аналитических функций 539
§ 16. Геометрический смысл аналитических функций 544
§ 17. Примеры конформных отображений 547

Алфавитный указатель 553

ПРЕДИСЛОВИЕ

Предлагаемая читателю книга третья «Энциклопедии элементарной математики» завершает первый большой раздел этого издания, посвящённый систематическому изложению тех элементов математической науки, на основе которых складываются школьные курсы арифметики, алгебры и отчасти тригонометрии. Если материал первых двух книг ограничивался преимущественно вопросами арифметики и алгебры в собственном смысле слова как учения о числах, их обобщениях, операциях над ними (имеются в виду алгебраические операции: сложение, вычитание, умножение и деление) и алгебраических уравнениях, то третья книга посвящена вопросам анализа, а именно, функциям и пределам. Наряду с учением об элементарных функциях и обстоятельно изложенной теорией пределов, сюда вошли также наиболее элементарные сведения из дифференциального и интегрального исчисления, теории рядов и сведения о функциях комплексного переменного.

Понятия производной и интеграла давно стучатся в двери общеобразовательной школы; как бы ни относиться к вопросу об их фактическом включении в школьные программы, сколько-нибудь удовлетворительное завершённое изложение элементарных основ математической науки без этих основных понятий следует признать немыслимым при современном состоянии науки.

Что касается функций комплексного переменного, то нет ни возможности, ни необходимости в том, чтобы вводить, хотя бы и не в близком будущем, систематические сведения о них в школьную программу. Однако тот основной факт, что элементарные функции являются аналитическими функциями, определёнными во всей комплексной плоскости (за исключением, быть может, определённых точек) и что, следовательно, полного понимания свойств этих функций и связей между ними можно достичь, только рассматривая их как функции комплексного переменного, оправдывает включение в нашу книгу небольшого очерка об аналитических функциях комплексного переменного.

Редакция

В. Л. ГОНЧАРОВ

ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ ДЕЙСТВИТЕЛЬНОГО ПЕРЕМЕННОГО. ПРЕДЕЛЫ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И ФУНКЦИЙ. ОБЩЕЕ ПОНЯТИЕ ФУНКЦИИ

ГЛАВА I

ОБЩИЕ СВЕДЕНИЯ ОБ ЭЛЕМЕНТАРНЫХ ФУНКЦИЯХ И ГРАФИКАХ УРАВНЕНИЙ

§ 1. Элементарные функции

Ближайшим предметом рассмотрения в математике являются числа и выполняемые над ними операции (действия). И понятие числа и понятие операции допускают неограниченные расширения и обобщения. В настоящей статье, если не сделано особой оговорки, речь будет итти лишь о действительных числах и (в трёх первых главах) преимущественно о тех операциях, которые изучаются в элементарной математике и потому сами носят название *элементарных*. Сюда относятся прежде всего алгебраические операции — сложение, вычитание, умножение и деление, затем возведение в произвольную степень и извлечение корня произвольной степени, логарифмирование по произвольному положительному основанию и, наконец, составление из данной величины шести тригонометрических функций (синус, косинус, тангенс, косеканс, секанс и котангенс), а также так называемых обратных функций (арксинус, арккосинус и т. д.).

Перечисленные операции могут выполняться, смотря по обстоятельствам, в том или ином заранее указанном порядке над данными числами или над буквами (переменными величинами), обозначающими числа. Мы будем предполагать, по крайней мере в пределах глав I и II, что число выполняемых операций конечно. Результат вычисления можно обозначить какой-нибудь новой буквой; при этом те буквы, которые участвовали в вычислении, ставятся в скобках, в определённом порядке, будучи разделены запятыми. Например:

$$v = u^{\operatorname{tg} u}, \tag{1}$$

или

$$f(t) = \frac{2t}{t^2 + 1}, \tag{2}$$

или ещё

$$F(x, y) = x^2 + y^2 - 25. \tag{3}$$

Получаемые таким образом математические выражения, или формулы, способные принимать то или иное числовое значение в зави-

симости от числовых значений входящих (т. е. участвующих в вычислении) величин, являются элементарными функциями этих величин.

Входящие в данную формулу «переменные» величины называются *независимыми*, сама же функция носит название *зависимой* переменной.

Функциональные символы вроде $f(t)$ или $F(x, y)$ особенно удобны в следующем отношении: что бы ни представляли собой величины c, a, b — числа или же новые буквенные выражения, — через $f(c)$ или $F(a, b)$ обозначают то, что получится, если вместо t подставить c, или вместо x подставить a, а вместо y подставить b. Так, из соотношений (2) и (3) вытекает:

$$f(3)=\frac{2\cdot 3}{3^2+1}=\frac{3}{5}, \qquad f(1)=\frac{2\cdot 1}{1^2+1}=1;$$

$$F(6,\ 7)=6^2+7^2-25=60, \quad F(4,\ 3)=4^2+3^2-25=0,$$

и точно так же

$$f\left(\frac{m}{n}\right)=\frac{2\cdot\frac{m}{n}}{\left(\frac{m}{n}\right)^2+1}=\frac{2mn}{m^2+n^2}, \quad f(-t)=\frac{2(-t)}{(-t)^2+1}=-\frac{2t}{t^2+1},$$

$$F(px,\ qy)=(px)^2+(qy)^2-25=p^2x^2+q^2y^2-25.$$

Если независимые переменные не выписаны в скобках (что имеет своё преимущество краткости), то в случае подстановок запись приходится усложнять; например, соотношение (1) даёт:

$$[v]_{u=\pi}=\pi^{\operatorname{tg}\pi}=1,$$

или же пользуются описательными оборотами речи: «при $u=\pi$ величина v принимает значение 1».

Читателю, несомненно, знакомо определение функции (от одной переменной) как соответствия между числовыми значениями независимой переменной и числовыми значениями зависимой переменной.

Переменная величина y называется функцией переменной величины x (в некотором промежутке I), *если каждому значению x* (из I) *соответствует некоторое определённое значение y.*

В таком общем виде определение функции было дано Н. И. Лобачевским в 1834 г. в следующих словах[1]):

«Это общее понятие (функции. — *Ред.*) требует, чтобы функцией от x называть число, которое даётся для каждого x и вместе с x постепенно изменяется. Значение функции может быть дано или аналитическим выражением или условием, которое подаёт средство испытывать все числа и выбирать одно из них; или, наконец, зависимость может существовать и оставаться неизвестной».

Слово «соответствует» (или «сопоставляется» иногда также говорят «отвечает») оставляет открытым вопрос о том, какова должна быть природа

---

[1]) Н. И. Лобачевский, Полное собрание сочинений, т. V, Гостехиздат, 1951, стр. 43.

правила, посредством которого устанавливается соответствие: важно лишь, чтобы такого рода правило было указано.

В частности, правило соответствия может иметь «эмпирический» характер; так, если говорят о температуре как функции времени, то правило заключается в том, чтобы в назначенный момент времени зафиксировать показание термометра.

В преподавании элементарной математики имеют особенно важное, если не исключительное, значение такие функции, для которых правило соответствия носит «оперативный» или «аналитический» характер: оно указывает, в надлежащем порядке, те математические действия (операции), которые надо совершить над значением x, чтобы получить значение y.

Нет оснований противопоставлять понятие однозначного аналитического выражения понятию функции как соответствия: первое является частным случаем второго [1]).

Понятие функции как аналитического выражения сложилось в первой половине XVIII в. Именно так определяли функцию И. Бернулли (1718 г.) и Л. Эйлер (1748 г.). Последний предложил следующее определение:

«Функция переменного количества есть аналитическое выражение, составленное каким-либо образом из этого переменного количества и чисел или постоянных количеств».

Следует, однако, заметить, что у Эйлера

1) не вполне чётко отграничены «допустимые» операции,

2) не исключаются формулы, содержащие бесчисленное множество операций.

Точное определение элементарной функции (в современном смысле) опирается на понятие функции как соответствия и формулируется так:

Функция называется элементарной, если её значения могут быть получены из постоянных чисел и значений независимых переменных посредством конечного числа элементарных операций.

Конкретные примеры неэлементарных функций приведены в главе IV. Там же указан и наиболее естественный способ их получения (см. § 49).

К понятию функции как соответствия нам придётся обращаться в данной статье неоднократно. Покуда же просим читателя, если идёт речь о «функциях», иметь в виду те самые элементарные функции, с которыми приходится встречаться в процессе преподавания.

В дальнейшем (в главах I—IV) число рассматриваемых переменных величин ограничивается двумя: ради единообразия они будут обозначены буквами x и y.

Пусть дано уравнение вида

$$F(x, y) = 0, \tag{4}$$

где $F(x, y)$ — какая-нибудь элементарная функция величин x и y [2]). Предположим, что x_0 и y_0 — произвольные числа. Если эти числа

---

[1]) Два понятия функции (более узкое и более широкое) можно сблизить между собой, или даже отождествить, одним из следующих способов:

а) устанавливая, что весьма обширные классы функций-соответствий допускают аналитическое представление (см., например, теорему Вейерштрасса в § 49 гл. IV);

б) рассматривая как математическую операцию переход от числового значения независимой переменной к соответствующему (в силу функционального соотношения) значению зависимой переменной.

[2]) Если в правой части уравнения стоит не нуль, всегда можно перенести всё в левую часть.

таковы, что при подстановке x_0 вместо x и y_0 вместо y [1]) уравнение удовлетворяется, т. е. его левая часть становится равной нулю,

$$F(x_0, y_0) = 0,$$

то пара чисел (x_0, y_0) представляет собой решение (одно из решений) данного уравнения. Функция $F(x, y)$ может оказаться такой, что уравнение $F(x, y) = 0$ не имеет вовсе решений $\left(\text{например, при } F(x, y) = \frac{1}{x+y} \text{ или при } F(x, y) = 2^{x-y}\right)$; или она может быть такая, что существует только одно решение или, вообще, конечное число решений (например, при $F(x, y) = x^2 + y^2$ имеется единственное решение: $x = 0$, $y = 0$); не исключена и «противоположная крайность», когда функция $F(x, y)$ обращается в нуль тождественно, так что любая пара значений x и y оказывается решением.

Более важным и часто встречающимся является иной случай, когда существует бесконечное множество решений уравнения $F(x, y) = 0$, и дело обстоит именно таким образом, что, задав «произвольно» значение какой-нибудь одной переменной, можно подыскать одно значение (или несколько) другой переменной так, чтобы уравнение удовлетворялось. Тогда говорят, что данное уравнение устанавливает функциональную зависимость между переменными x и y.

Предположим, например, что

$$F(x, y) = 2x - 5y + 10. \tag{5}$$

Уравнение

$$2x - 5y + 10 = 0 \tag{6}$$

таково, что, задав значение x совершенно произвольно, можно найти решение, если взять значение y согласно формуле

$$y = \frac{2x + 10}{5}. \tag{7}$$

Таким образом, каждому значению x соответствует одно определённое значение y, удовлетворяющее нашему уравнению: оно даётся предыдущей формулой. Если положить

$$f(x) = \frac{2x + 10}{5},$$

то можно сказать, что уравнение $F(x, y) = 0$ равносильно уравнению $y = f(x)$.

---

[1]) Предполагается, что эта подстановка «имеет смысл», т. е. что совокупность операций, указываемых символом F, может быть выполнена при значениях $x = x_0$, $y = y_0$.

В качестве второго примера рассмотрим функцию $F(x, y) = x^2 + y^2 - 1$. Свойства уравнения

$$x^2 + y^2 - 1 = 0 \tag{8}$$

иные: здесь можно произвольно взять значение x лишь из промежутка $-1 \leqslant x \leqslant +1$, и тогда получаются д в а различных значения y:

$$y = \pm \sqrt{1 - x^2}.$$

Положив

$$f_1(x) = +\sqrt{1 - x^2}, \quad f_2(x) = -\sqrt{1 - x^2},$$

легко понять, что каждое из уравнений

$$y = f_1(x), \quad y = f_2(x) \tag{9}$$

влечёт за собой уравнение (8), тогда как из уравнения (8) при любом значении x из рассматриваемого промежутка следует или одно или другое из уравнений (9) [1]).

Роли переменных x и y в этом примере можно было бы поменять.

Возвращаясь от частных примеров к общему случаю, следует сказать, что имеется существенное преимущество в том, чтобы рассматривать функциональную зависимость между x и y в виде уравнения, связывающего между собой переменные x и y, тем самым оставляя за собой право, если угодно, считать x независимой переменной, а y — зависимой, или наоборот.

Предположим, например, что в качестве независимой переменной мы хотим взять величину x. Если случится, что, каково бы ни было выбранное значение x из некоторого промежутка (например, при $a < x < b$), уравнение $F(x, y) = 0$ имеет всегда один корень относительно y, и этот корень удастся выразить в виде элементарной формулы, зависящей, естественно, от x, тогда величина y, удовлетворяющая данному уравнению, в рассматриваемом промежутке является о д н о з н а ч н о й ф у н к ц и е й величины x:

$$y = f(x).$$

Последнее уравнение в этом случае равносильно заданному уравнению $F(x, y) = 0$ (в рассматриваемом промежутке). Но может случиться и так, что каждому значению x из некоторого промежутка соответствует таким же образом несколько (например, k) корней (относительно y) данного уравнения; предполагая, что каждый из этих k корней определяется по особой элементарной формуле

$$y = f_1(x), \quad y = f_2(x), \quad \ldots, \quad y = f_k(x),$$

мы будем иметь в этом случае k различных однозначных функций, или, как иногда говорят, одну м н о г о з н а ч н у ю (k-з н а ч н у ю)

[1]) К уравнению (8) мы вернёмся в § 46, см. стр. 213.

функцию в рассматриваемом промежутке. В этом случае заданное уравнение $F(x, y)=0$ «расщепляется» на k уравнений.

В первом из приведённых выше примеров уравнение (6) определяет y как однозначную функцию x для каких угодно значений x $(-\infty<x<+\infty)$; во втором — уравнение (8) определяет y как двузначную функцию x в промежутке $-1\leqslant x\leqslant+1$.

Чтобы убедиться в том, насколько разнообразны возникающие здесь возможности, рассмотрим ещё третий пример [1]), полагая

$$F(x, y)=x^4+y^4-x^2-y^2. \tag{10}$$

Уравнение

$$x^4+y^4=x^2+y^2 \tag{11}$$

(что равносильно $x^4+y^4-x^2-y^2=0$) — биквадратное относительно y, и решения его даются формулами

$$y=\pm\sqrt{\frac{1}{2}\pm\sqrt{\frac{1}{4}+x^2-x^4}}.$$

Легко убедиться, что при

$$1<|x|<\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}$$

(т. е. в промежутках

$$-\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}<x<-1 \quad \text{и} \quad 1<x<\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}\Big)$$

выражение, стоящее под внутренним радикалом, положительно, и сам этот радикал меньше чем $\frac{1}{2}$, так что выражение, стоящее под внешним радикалом, независимо от выбора знака под радикалом, будет также положительным; таким образом, в указанных промежутках уравнение (11) приводит нас к рассмотрению четырёх однозначных элементарных функций

$$y=f_i(x) \qquad (i=1,\ 2,\ 3,\ 4),$$

причём положено

$$f_1(x)=-\sqrt{\frac{1}{2}+\sqrt{\frac{1}{4}+x^2-x^4}},$$

$$f_2(x)=-\sqrt{\frac{1}{2}-\sqrt{\frac{1}{4}+x^2-x^4}},$$

$$f_3(x)=+\sqrt{\frac{1}{2}-\sqrt{\frac{1}{4}+x^2-x^4}},$$

$$f_4(x)=+\sqrt{\frac{1}{2}+\sqrt{\frac{1}{4}+x^2-x^4}}.$$

[1]) См. стр. 76.

В промежутке $-1<x<+1$ выражение, стоящее под внутренним радикалом, положительно, но сам внутренний радикал больше чем $\frac{1}{2}$, и потому в этом промежутке имеется л и ш ь д в а значения:

$$y=f_1(x) \quad \text{и} \quad y=f_2(x).$$

Наконец, при выполнении условия

$$|x|>\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}$$

(т. е. в промежутках

$$-\infty<x<-\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}} \quad \text{и} \quad \sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}<x<\infty)$$

уравнение (11) не имеет вовсе корней.

§ 2. Графические представления. Приёмы точечных построений

Мы убедились, что уравнение вида

$$F(x, y)=0 \tag{12}$$

может иметь с к о л ь к о у г о д н о решений.

Чтобы придать совокупности решений бо́льшую обозримость, прибегают к плоскости Oxy, и с каждым решением (x, y) сопоставляют точку с абсциссой x и ординатой y. Все получаемые таким образом точки-решения, будучи рассматриваемы как целое, образуют г р а ф и к уравнения. Точнее и проще: *графиком уравнения называется совокупность* (множество) *всех точек* [1]) *плоскости, координаты которых удовлетворяют уравнению.*

Нахождение всех решений данного уравнения и построение его графика — в сущности равносильные задачи, но, говоря о графике уравнения, а не о совокупности его решений, мы не только придаём наглядность интересующему нас вопросу, но и значительно упрощаем речь.

В случае, если данное уравнение «решено» относительно зависимой переменной y, так что эта переменная y в некотором промежутке $a<x<b$ представляется в виде однозначной функции переменной x

$$y=f(x),$$

то говорят без всякого различия о «графике у р а в н е н и я $y=f(x)$», или о «графике ф у н к ц и и $f(x)$».

---

[1]) Можно также сказать: «геометрическое место точек»...

Характерное свойство графика в этом случае то, что всякая прямая, параллельная оси Oy (в пределах промежутка), имеет с графиком ровно одну общую точку.

Указанный случай является особенно важным как практически, так и теоретически; к нему в дальнейшем преимущественно направлено внимание.

Чтобы построить график данного уравнения, нужно, вообще говоря, отметить в плоскости Oxy все принадлежащие ему точки. Этого сделать на самом деле, конечно, нельзя (если не говорить о частных случаях) по той простой причине, что точек графика — бесконечное множество. Обыкновенно делают приближённое построение и именно следующим образом: отмечают в плоскости достаточное количество точек, стараясь вместе с тем, чтобы последовательно отмечаемые точки находились одна от другой на достаточно близких расстояниях; соединяя затем последовательно отмеченные точки «плавной линией», получают график уравнения. Такая процедура носит название *построения по точкам*. При этом, конечно, в зависимости от характера производимых вычислений и геометрических операций, отмечаемые на чертеже точки бывают определены более или менее точно; но насколько точно воспроизведён бывает график в промежутках «между» отмеченными точками,

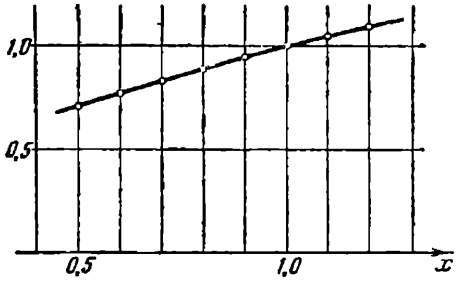


Рис. 1.

x	y	x	y
0,5	0,71	0,9	0,95
0,6	0,77	1,0	1,00
0,7	0,84	1,1	1,05
0,8	0,89	1,2	1,10

это зависит, с одной стороны, от свойств данного уравнения, с другой — от знаний и опытности производящего построение.

Общеизвестный вычислительный приём построения графиков функций заключается в том, что для ряда значений независимой переменной x [1]) вычисляют значения функции $y=f(x)$, записывают результаты в легко обозримой табличной форме, вслед за тем (или параллельно с вычислениями) отмечают соответствующие точки на чертеже и соединяют их плавной кривой.

---

[1]) Часто эти значения берут «равноотстоящими», т. е. образующими арифметическую прогрессию. Если, например, берутся значения, выражающиеся десятичными дробями с одним знаком после запятой, то говорят для краткости, что берутся значения «через одну десятую».

В качестве примера приведём таблицу и рис. 1, составленные для функции

$$y=\sqrt{x} \tag{13}$$

в промежутке $0{,}5 \leqslant x \leqslant 1{,}2$, причём значения x взяты через одну десятую.

В следующем примере (рис. 2) в таблице, составленной для функции

$$y=\frac{x(x+1)}{x+2}, \tag{14}$$

значения x взяты через одну пятую, в промежутке $0 \leqslant x \leqslant 1$.

x	y	x	y
0	0 или 0,00	$\frac{3}{5}$	$\frac{24}{65}$ или 0,37
$\frac{1}{5}$	$\frac{6}{55}$ » 0,11	$\frac{4}{5}$	$\frac{18}{35}$ » 0,51
$\frac{2}{5}$	$\frac{7}{30}$ » 0,23	1	$\frac{2}{3}$ » 0,67

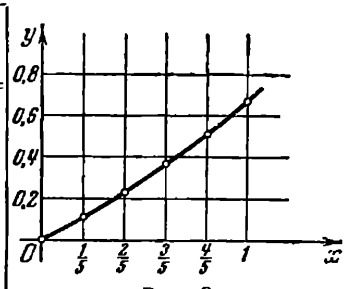


Рис. 2.

В таблице иногда бывает удобно вставлять промежуточные столбцы, заполняя их последовательно. Это видно на примере таблицы, составленной для функции:

$$y=\frac{1}{2}\sqrt{x^2+1}. \tag{15}$$

x	x^2	x^2+1	$\sqrt{x^2+1}$	$\frac{1}{2}\sqrt{x^2+1}$
0,0	0,00	1,00	1,00	0,50
0,5	0,25	1,25	1,12	0,56
1,0	1,00	2,00	1,41	0,70
1,5	2,25	3,25	1,80	0,90
2,0	4,00	5,00	2,24	1,12
2,5	6,25	7,25	2,69	1,34

Наряду с вычислительным приёмом построения заслуживает внимания и геометрический. Предположим, что правая часть уравнения

$$y=f(x)$$

не содержит иных операций, кроме четырёх арифметических и извле-

чения квадратного корня; тогда, считая значение x данным в виде отрезка, можно построить отрезок y с помощью циркуля и линейки.

Если требуется сделать несколько подстановок

$$x = x_1, \quad x = x_2, \quad \ldots, \quad x = x_n,$$

то целесообразно, конечно, систематизировать работу и производить n построений совместно, последовательно выполняя один и тот же этап во всех n построениях.

Предположим, например, что дана та же функция (15). Чтобы произвести отдельное построение, возьмём точку P с абсциссой x на горизонтальной [1]) оси и точку Q с ординатой 1 на вертикальной оси (рис. 3); через точку P проведём вертикальную прямую и на ней отложим вверх отрезок PR, равный PQ; если разделим затем отрезок PR пополам, то точка деления M будет иметь как раз абсциссу x и ординату y.

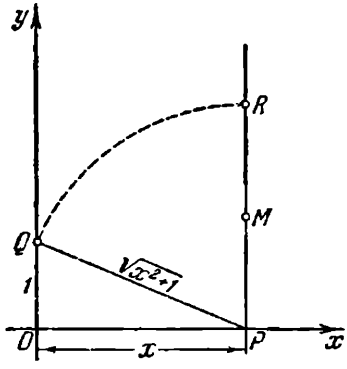


Рис. 3.

Пусть требуется построить те точки графика (15), которые соответствуют абсциссам $x_1, x_2, \ldots, x_n$. Тогда: 1) на горизонтальной оси мы отметим точки $P_1, P_2, \ldots, P_n$ с абсциссами $x_1, x_2, \ldots, x_n$, а на вертикальной оси — точку Q с ординатой 1; 2) через точки $P_1, P_2, \ldots, P_n$ проведём вертикальные прямые; 3) посредством циркуля на прямой, проходящей через P_1, отложим отрезок P_1R_1, равный P_1Q; затем на прямой, проходящей через P_2, отложим отрезок P_2R_2, равный P_2Q, и т. д.; 4) найдём середину M_1 отрезка P_1R_1, середину M_2 отрезка P_2R_2 и т. д.

Точки $M_1, M_2, \ldots$, очевидно, и будут искомыми.

В другом примере

$$y = \frac{x(x+1)}{x+2} \qquad (16)$$

построение можно выполнить следующим образом.

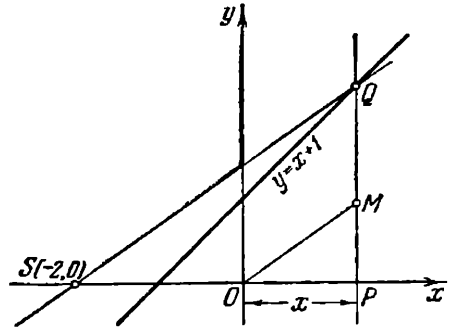


Рис. 4.

Отметим в плоскости Oxy точку S с координатами $(-2, 0)$ и проведём прямую $y = x + 1$, отсекающую на осях Ox и Oy соответственно отрезки -1 и 1 (рис. 4). Возьмём на горизонтальной оси точку P с абсциссой x, проведём

[1]) В дальнейшем мы для удобства речи называем ось Ox горизонтальной, ось Oy — вертикальной, а прямые, параллельные той или иной оси, — горизонтальными и вертикальными.

через неё вертикальную прямую и отметим точку пересечения Q с прямой $y=x+1$.

В треугольнике PQS катеты PQ и PS (при $x>-1$) соответственно равны $x+1$ и $x+2$. Поэтому достаточно из начала координат O провести прямую, параллельную SQ до пересечения с PQ в точке M, чтобы получить треугольник PMO, подобный PQS; из подобия же следует, что

$$\frac{PM}{OP}=\frac{PQ}{SP},$$

т. е.

$$PM=\frac{OP\cdot PQ}{SP}=\frac{x(x+1)}{x+2},$$

так что точка M как раз принадлежит графику.

Чтобы получить несколько точек графика M_1, M_2, ..., M_n с абсциссами x_1, x_2, ..., x_n, достаточно: 1) отметить на горизонтальной оси точки P_1, P_2, ..., P_n, с такими же абсциссами, 2) провести через них вертикали до пересечения с прямой $y=x+1$ в точках Q_1, Q_2, ..., Q_n, 3) через точку S провести прямые SQ_1, SQ_2, ..., SQ_n, 4) через начало координат O провести прямые, параллельные SQ_1, SQ_2, ..., SQ_n, до пересечения с прямыми P_1Q_1, P_2Q_2, ..., P_nQ_n. Точки пересечения M_1, M_2, ..., M_n будут искомыми.

Рассмотрим, наконец, пример

$$y=\sqrt{x}. \tag{17}$$

Руководствуясь тем, что величину y можно понимать как среднюю геометрическую из величин x и 1, отметим на горизонтальной оси точки S с абсциссой -1 и P с абсциссой x. На отрезке SP как на диаметре построим окружность и из точки N её пересечения с вертикальной осью опустим перпендикуляр на вертикальную прямую, проведённую через P. Обозначая основание перпендикуляра буквой M, видим, что точка M имеет абсциссу x и ординату

$$PM=ON=\sqrt{x}$$

и, следовательно, принадлежит графику (рис. 5).

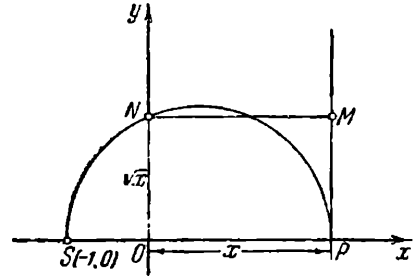


Рис. 5.

Таким образом, точечное построение сводится к проведению через точку S ряда окружностей с центрами на горизонтальной оси и к проектированию точек пересечения окружностей с вертикальной осью на касательную, проведённую через противоположный конец диаметра.

Примеры такого рода можно разнообразить, изощряя изобретательность в геометрических построениях.

Рассмотрим следующее вполне естественное обобщение. Допустим, что заданы, в качестве вспомогательного средства построения, графики функций

$$y=u(x),\quad y=v(x),\ldots$$

В таком случае можно построить по точкам график функции

$$y=f(x,\ u(x),\ v(x),\ldots), \tag{18}$$

если только допустить, что правая часть уравнения составляется из своих аргументов x, u, v, ... посредством пяти упомянутых выше операций. Действительно, для каждого заданного отрезка x известными являются, по предположению, отрезки $u(x)$, $v(x)$, ..., и из них с помощью циркуля и линейки может быть построен отрезок $f(x,\ u(x),\ v(x),\ldots)$.

Нас в особенности интересуют следующие частные случаи графических построений:

1. Если заданы графики функций $y=u(x)$ и $y=v(x)$, то можно построить по точкам график функции

$$y=u(x)+v(x) \tag{19}$$

(рис. 6, *а*).

2. Если заданы графики функций $y=u(x)$ и $y=v(x)$, то можно построить по точкам график функции

$$y=u(x)-v(x). \tag{20}$$

3. Если заданы графики функций $y=u(x)$ и $y=v(x)$, то можно построить по точкам график функции

$$y=u(x)\ v(x) \tag{21}$$

(рис. 6, *б*).

4. Если заданы графики функций $y=u(x)$ и $y=v(x)$, то можно построить по точкам график функции

$$y=\frac{u(x)}{v(x)}. \tag{22}$$

Легко понять, что дело сводится к сложению, вычитанию, умножению и делению отрезков. «Отрезки» понимаются, конечно, в алгебраическом смысле, так как $u(x)$ и $v(x)$ представляют собой действительные числа, которые могут быть больше нуля, меньше нуля или равны нулю. «Умножение» понимается в том смысле, что перемножены должны быть числа и взят затем отрезок

(ордината точки), равный по длине и по знаку полученному произведению; аналогично — для деления [1]).

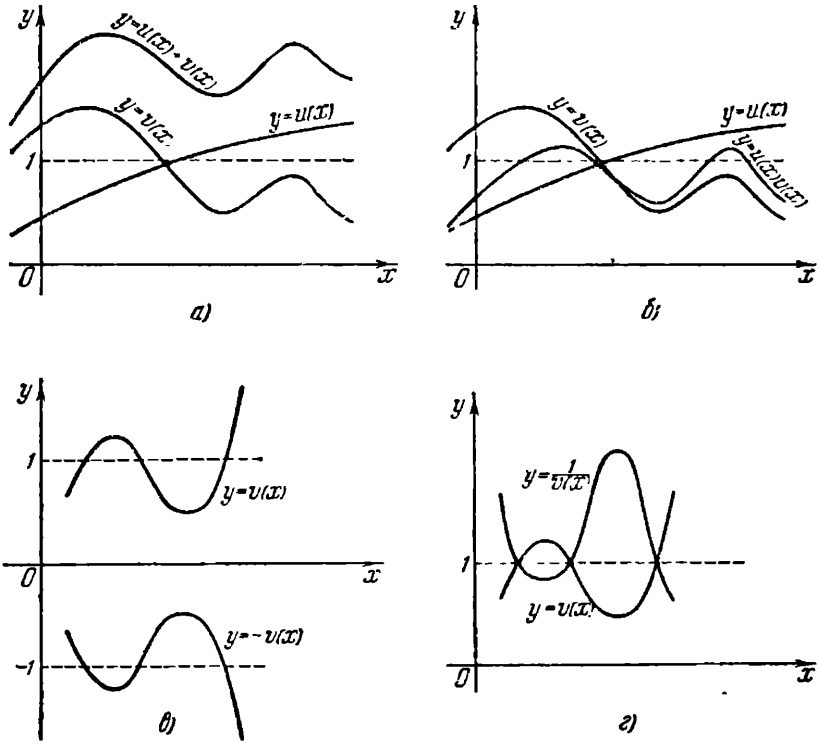


Рис. 6.

Обыкновенно бывает целесообразно избегать вычитания и деления, сводя вычитание к сложению, а деление к умножению:

$$u - v = u + (-v),$$

$$\frac{u}{v} = u \cdot \frac{1}{v},$$

и опираясь при этом на следующие положения:

5. Если задан график функции $y = v(x)$, то не представляет труда получить и график функции $y = -v(x)$ (рис. 6, *в*).

6. Если задан график функции $y = v(x)$, то можно легко построить по точкам и график функции $y = \frac{1}{v(x)}$ (рис. 6, *г*).

---

[1]) Геометрическое построение в случаях 3 и 4 сводится, очевидно, к нахождению четвёртой пропорциональной, так как равенству $y = uv$ можно придать вид $\frac{y}{u} = \frac{v}{1}$, а равенству $y = \frac{u}{v}$ — вид $\frac{y}{1} = \frac{u}{v}$.

В самом деле, из графика $y=v(x)$ график $y=-v(x)$ получается посредством замены всех ординат им противоположными по знаку, т. е. посредством симметрического отражения относительно оси Ox; что же касается графика $y=\frac{1}{v(x)}$, то он получается из графика $y=v(x)$ посредством замены всех ординат им обратными.

Приёмы построений, вытекающие из пунктов 1—4, можно назвать кратко «сложением», «вычитанием», «умножением» и «делением» графиков. Выполнять каждую из этих операций можно путём вычисления или путём геометрических построений.

В данной связи не излишне указать ещё и на следующее положение, не являющееся частным случаем раньше сформулированного общего принципа:

7. Если заданы графики функций $y=u(x)$ и $y=v(x)$, то можно построить по точкам график «сложной функции» («функции от функции»):

$$y=u(v(x)). \tag{23}$$

Рассмотрим этот вопрос подробнее. Пусть P — точка с абсциссой x на горизонтальной оси (рис. 7). Проводя через точку P вертикальную прямую, в пересечении её с графиком $y=v(x)$ поставим точку Q; отрезок PQ равен $v(x)$. Пусть T — точка на горизонтальной оси с абсциссой $v(x)$ (так что $OT=PQ$); тогда, проводя через точку T вертикаль до пересечения с графиком $y=u(x)$ в точке S, получим отрезок TS, равный $u(v(x))$. Нам останется спроектировать точку S на вертикаль, проведённую через P, чтобы получить точку M с абсциссой x и ординатой $u(v(x))$.

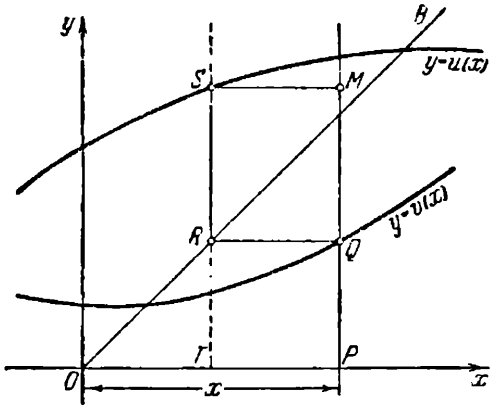


Рис. 7.

Вместо того, чтобы откладывать на горизонтальной оси отрезок OT, равный PQ, можно воспользоваться биссектрисой OB координатного угла Oxy (уравнение которой есть $x=y$): достаточно из точки Q провести горизонтальный отрезок до пересечения с OB в точке R, затем из точки R провести вертикальный отрезок RS до пересечения с графиком $y=u(x)$ в точке S; далее, как указано выше.

Разобравшись в этом построении, можно очень быстро выполнить его для целого ряда точек P, проводя сначала ряд отрезков PQ, затем ряд отрезков QR, затем ряд отрезков RS и, наконец, ряд отрезков SM (см. рис. 7).

§ 3. Простейшие преобразования графиков

Идея, лежащая в основе излагаемых ниже соображений, заключается в следующем. Предположим, что тем или иным путём уже построен график G уравнения

$$F(x, y) = 0. \tag{24}$$

Тогда в ряде случаев, если только уравнение $F_*(x, y) = 0$ отличается от уравнения $F(x, y) = 0$ определённым, легко устанавливаемым признаком, график G_* уравнения $F_* = 0$ получается из графика G уравнения $F = 0$ посредством некоторого легко выполнимого геометрического преобразования.

Мы ограничимся рассмотрением следующих элементарных преобразований:

I. а) Осевая симметрия относительно оси Ox.

б) Осевая симметрия относительно оси Oy.

в) Центральная симметрия относительно начала O.

(Симметрии относительно произвольной прямой и относительно произвольного центра рассматривать не будем.)

II. а) Перенесение (сдвиг, трансляция) на данный отрезок параллельно оси Ox.

б) То же параллельно оси Oy.

(Перенесения, параллельного произвольной прямой, рассматривать не будем.)

III. а) Растяжение (или сжатие) в данном отношении по направлению оси Ox[1]).

б) То же по направлению оси Oy.

(Растяжения по произвольному направлению рассматривать не будем.)

Иных преобразований, кроме перечисленных элементарных, мы рассматривать не будем; однако нам придётся встречаться с преобразованиями, возникающими как результат последовательного выполнения нескольких элементарных. Останутся в стороне также преобразования вращения (если не считать одного частного случая, см. ниже теорему IV).

Свойства перечисленных преобразований, которые могут служить их определениями, заключаются в следующем:

I. а) При осевой симметрии относительно оси Ox точка (x, y) переходит в точку $(x, -y)$.

I. б) При осевой симметрии относительно оси Oy точка (x, y) переходит в точку $(-x, y)$.

I. в) При центральной симметрии относительно начала координат O точка (x, y) переходит в точку $(-x, -y)$.

---

[1]) «Сжатие в m раз» — то же, что «растяжение в $\frac{1}{m}$ раз».

II. а) При перенесении на отрезок $a\,(a \gtrless 0)$ параллельно оси Ox точка (x, y) переходит в точку $(x+a, y)$.

II. б) При перенесении на отрезок $b\,(b \gtrless 0)$ параллельно оси Oy точка (x, y) переходит в точку $(x, y+b)$.

III. а) При растяжении в p раз $(p>0,\ p \gtrless 1)$ по направлению оси Ox точка (x, y) переходит в точку (px, y).

III. б) При растяжении в q раз $(q>0,\ q \gtrless 1)$ по направлению оси Oy точка (x, y) переходит в точку (x, qy).

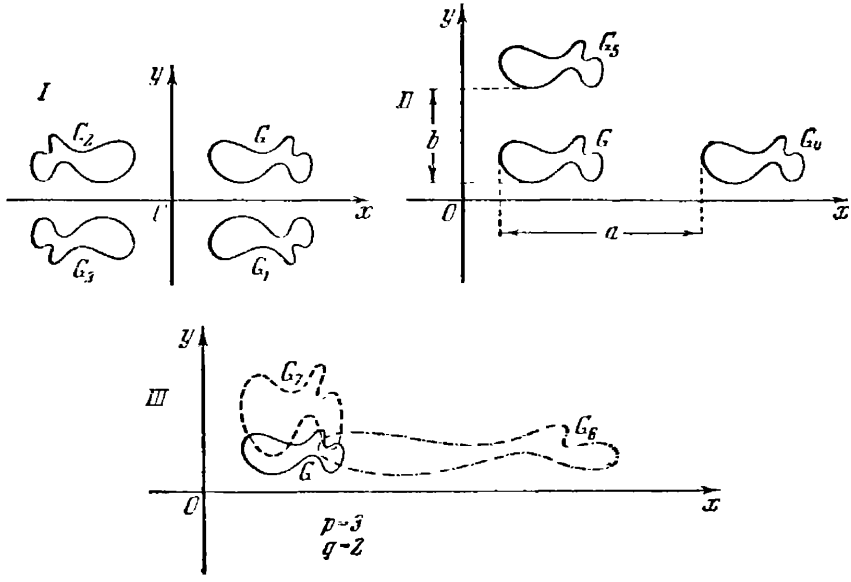


Рис. 8.

Теперь можно сказать, что справедливы следующие утверждения (рис. 8):

I. а) График G_1 уравнения

$$F(x, -y)=0$$

симметричен графику G уравнения $F(x, y)=0$ относительно оси Ox.

I. б) График G_2 уравнения

$$F(-x, y)=0$$

симметричен графику G уравнения $F(x, y)=0$ относительно оси Oy.

I. в) График G_3 уравнения

$$F(-x, -y)=0$$

симметричен графику G уравнения $F(x, y)=0$ относительно начала координат O.

II. а) График G_4 уравнения

$$F(x-a,\ y)=0$$

получается из графика G уравнения $F(x,\ y)=0$ посредством перенесения параллельно оси Ox на отрезок, равный a.

II. б) График G_5 уравнения

$$F(x,\ y-b)=0$$

получается из графика G уравнения $F(x,\ y)=0$ посредством перенесения параллельно оси Oy на отрезок, равный b.

III. а) График G_6 уравнения

$$F\left(\frac{x}{p},\ y\right)=0$$

получается из графика G уравнения $F(x,\ y)=0$ посредством растяжения в p раз по направлению оси Ox.

III. б) График G_7 уравнения

$$F\left(x,\ \frac{y}{q}\right)=0$$

получается из графика G уравнения $F(x,\ y)=0$ посредством растяжения в q раз по направлению оси Oy.

На рис. 8 *III* изображены график G_6 при $p=3$ и график G_7 при $q=2$.

Доказательство I. а) Пусть точка $M(x_0,\ y_0)$ принадлежит графику G, так что

$$F(x_0,\ y_0)=0.$$

В таком случае точка $M'(x_0,\ -y_0)$ принадлежит графику G_1, так как её координаты удовлетворяют уравнению $F(x,\ -y)=0$. Но точка M' симметрична точке M относительно оси Ox. Итак, график G_1 содержит все точки, симметричные какой-нибудь точке графика G относительно Ox. Если бы график G содержал ещё какую-нибудь лишнюю точку $N'(x_1,\ y_1)$, то имело бы место равенство

$$F(x_1,\ -y_1)=0,$$

и тогда вышло бы, что симметричная точка $N(x_1,\ -y_1)$ принадлежит графику G, что противоречит допущению.

Доказательство II. а) Пусть точка $M(x_0,\ y_0)$ принадлежит графику G, так что

$$F(x_0,\ y_0)=0.$$

В таком случае точка $M'(x_0+a,\ y_0)$ принадлежит графику G_4, так как её координаты удовлетворяют уравнению $F(x-a,\ y)=0$. Но точка M', очевидно, получается из точки M посредством перенесения параллельно оси Ox. И т. д.

Доказательство III. а) Пусть точка $M(x_0, y_0)$ принадлежит графику G, так что

$$F(x_0,\ y_0)=0.$$

В таком случае точка $M'(px_0, y_0)$ принадлежит графику G_Φ, так как её координаты удовлетворяют уравнению $F\left(\frac{x}{p}, y\right)=0$. Но точка M' получается из точки M посредством растяжения в p раз по направлению оси Ox. И т. д. Доказательства остальных теорем аналогичны.

Формулируем установленные выше теоремы применительно к графикам функций.

Предположим, что данное уравнение имеет вид

$$y=f(x). \tag{25}$$

В этом случае можно положить

$$F(x, y)=y-f(x).$$

Но тогда

$$F(x, -y)=-y-f(x),$$

и уравнение $F(x, -y)=0$ может быть записано в виде

$$y=-f(x).$$

Отсюда вытекает теорема

I'. а) График функции $-f(x)$ симметричен графику функции $f(x)$ относительно оси Ox (см. § 2, п° 5).

Подобным же образом получаются и следующие теоремы:

I'. б) График функции $f(-x)$ симметричен графику функции $f(x)$ относительно оси Oy.

I'. в) График функции $-f(-x)$ симметричен графику функции $f(x)$ относительно начала координат O.

II'. а) График функции $f(x-a)$ получается из графика функции $f(x)$ посредством перенесения параллельно оси Ox на отрезок a.

II'. б) График функции $f(x)+b$ получается из графика функции $f(x)$ посредством перенесения параллельно оси Oy на отрезок b.

III'. а) График функции $f\left(\frac{x}{p}\right)$ получается из графика функции $f(x)$ посредством растяжения в p раз по направлению оси Ox.

III'. б) График функции $qf(x)$ получается из графика функции $f(x)$ посредством растяжения в q раз по направлению оси Oy.

Рассмотренные нами преобразования, как уже было замечено, могут комбинироваться между собой. Так, заменяя в общем уравнении $F(x, y)=0$ сначала x через $x-a$, затем y через $y-b$ (или в обратном порядке), получим уравнение

$$F(x-a, y-b)=0,$$

график которого получается из графика данного уравнения посредством перенесения на отрезок a параллельно оси Ox и на отрезок b параллельно оси Oy (что, конечно, равносильно одному перенесению параллельно некоторой наклонной прямой).

Или ещё: график уравнения

$$F\left(\frac{x}{p}, \frac{y}{q}\right)=0$$

получается из графика уравнения $F(x, y)=0$ посредством растяжения в p раз по направлению оси Ox и растяжения в q раз по направлению оси Oy.

Подобным же образом график уравнения

$$F(-x, -y)=0$$

получается из графика $F(x, y)=0$ посредством двух последовательных преобразований симметрии — сначала относительно оси Ox, затем относительно оси Oy (или наоборот). Но он же (согласно Iв) получается из того же графика посредством преобразования центральной симметрии относительно начала координат. Это не удивительно: легко показать и непосредственно геометрически, что последовательные симметричные отражения данной точки относительно двух взаимно перпендикулярных прямых дают в итоге точку, симметричную заданной относительно точки пересечения этих прямых.

Следует обратить особое внимание на тот случай, когда функция F не изменяется при замене y на $-y$:

$$F(x, -y)\equiv F(x, y).$$

В этом случае графики двух уравнений

$$F(x, -y)=0 \quad \text{и} \quad F(x, y)=0$$

совпадают и так как (вследствие Ia) они взаимно симметричны относительно оси Ox, то выписанное выше тождество свидетельствует о том, что график G уравнения $F(x, y)=0$ имеет ось Ox своей осью симметрии. Проведя аналогичные рассуждения с использованием Iб и Iв, мы можем составить следующую таблицу признаков наличия элементов симметрии:

если	то
$F(x, -y)\equiv F(x, y)$,	G имеет ось симметрии Ox
$F(-x, y)\equiv F(x, y)$,	G имеет ось симметрии Oy
$F(-x, -y)\equiv F(x, y)$,	G имеет центр симметрии O

Обратимся к случаю уравнения вида $y=f(x)$. Если функция $f(x)$ не изменяется при замене x на $-x$:

$$f(-x)\equiv f(x),$$

то её называют *чётной*; если при замене x на $-x$ абсолютная величина $f(x)$ не изменяется, но изменяется знак:

$$f(-x)\equiv -f(x),$$

то функцию называют *нечётной*.

Из теорем I'б и I'в следует:

если	то
функция $f(x)$ — чётная,	её график G имеет ось симметрии Oy
функция $f(x)$ — нечётная,	её график имеет центр симметрии O

Если функция $F(x, y)$ не изменяется при одновременной замене x на $x-a$ и y на $y-b$:

$$F(x-a, y-b) \equiv F(x, y),$$

то её называют *периодической с периодом* (a, b); то же в этом случае говорят и о графике уравнения $F(x, y) \equiv 0$. Если функция $f(x)$ не изменяется при замене x на $x-a$, то её (а также и её график) называют *периодической с периодом* a. На рис. 9, *а* и 9, *б* изображены соответственно графики функций $F(x, y)=0$ с периодом $(1, 2)$ и $y=f(x)$ с периодом, равным 1.

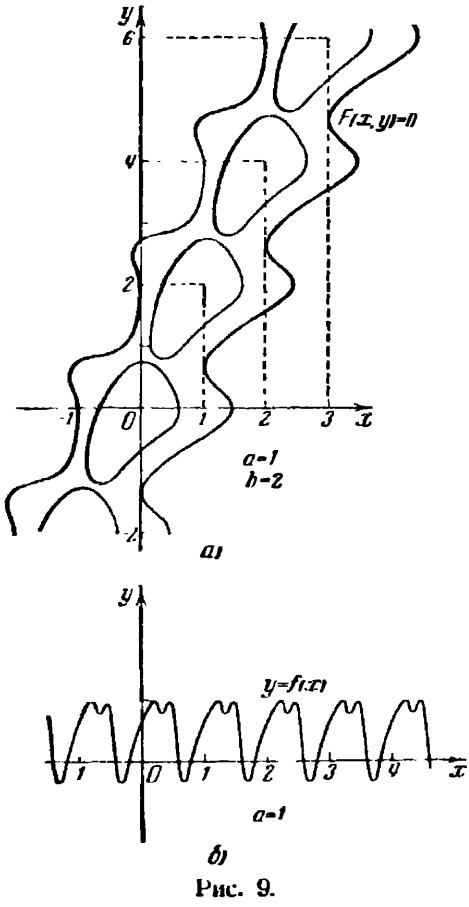


Рис. 9.

Если функция $f(x)$ имеет период a, то очевидно, что числа $2a$, $3a$, ..., na, ..., а также $-a$, $-2a$, ... в равной степени являются периодами. По большей части, говоря о периоде, имеют в виду наименьший из положительных периодов.

Подобные же соображения относятся и к функциям $F(x, y)$ и их графикам.

Можно было бы рассмотреть свойство функций не изменяться при замене x на px и y на qy

$$F(px, qy) \equiv F(x, y), \quad f(px) \equiv f(x),$$

чему соответствовало бы то обстоятельство, что график не изменяется при определённого рода растяжениях. Но это не представляет особенного интереса.

Установить сразу для данного графика наличие элементов симметрии или периодичности выгодно в том отношении, что позволяет

сократить последующую работу построения. Если установлено, что график имеет симметрию относительно вертикальной оси, то достаточно исследовать его лишь в одной (например, правой) полуплоскости; если установлено, что имеется симметрия относительно обеих осей, то можно ограничиться одним (например, первым) квадрантом; если известно, что график имеет период a, то достаточно рассмотреть его расположение в пределах одной «полосы периодов» $\left(\text{например, в полосе } 0 \leqslant x < a \text{ или } -\frac{a}{2} < x \leqslant \frac{a}{2}\right)$; и т. п.

Рассмотрим, ради разъяснения предшествующего, несколько конкретных примеров.

Пример 1. График уравнения $y = 4x^2$ имеет вертикальную ось симметрии. Переписав уравнение в виде $\frac{y}{4} = x^2$, видим, что из графика $y = x^2$ он получается посредством растяжения в 4 раза по направлению оси Oy. Написав то же уравнение в виде $y = \left(\frac{x}{1/2}\right)^2$, видим, что тот же график можно получить из графика $y = x^2$ посредством сжатия в 2 раза по направлению оси Ox.

Пример 2. График $y = 2x - 1$ получается из графика $y = 2x$ посредством перенесения параллельно оси Oy на отрезок -1. Написав уравнение в виде $y = 2\left(x - \frac{1}{2}\right)$, видим, что он может быть получен из того же графика посредством перенесения параллельно оси Ox на отрезок $\frac{1}{2}$.

Пример 3. Уравнение $x^2 + y^2 = 2(3x + 4y)$ можно написать в виде

$$(x - 3)^2 + (y - 4)^2 = 25$$

(«выделение квадратов») и тогда видно, что его график получается из графика $x^2 + y^2 = 25$ посредством перенесения по направлению оси Ox на отрезок 3 и по направлению оси Oy на отрезок 4.

Пример 4. График уравнения $y = 2^{x-3}$ получается из графика $y = 2^x$ посредством перенесения параллельно оси Ox на отрезок 3. Придав уравнению вид $y = \frac{1}{8} \cdot 2^x$, убеждаемся, что он получается из того же графика посредством сжатия в 8 раз по направлению оси Oy.

Пример 5. Функции $y = \cos x$, $y = \sin 2x$, $y = \cos x + \sin 2x$ имеют соответственно период 2π, π и 2π. Первая из них — чётная, так что график симметричен относительно оси Oy, вторая — нечётная, так что график имеет центр симметрии в начале O. Третья не является ни чётной, ни нечётной.

Пример 6. Обе функции $y = \operatorname{tg} \pi x$ (нечётная) и $y = \operatorname{tg}^2 \pi x$ (чётная) имеют один и тот же период 1.

Пример 7. График уравнения $\sin^2 x + \sin^2 y = \frac{1}{2}$ симметричен относительно обеих осей Ox и Oy и (следовательно) относительно центра O. Он имеет периоды $(\pi, 0)$ и $(0, \pi)$ (и вообще — периоды вида $(m\pi, n\pi)$, где m и n — целые числа).

В заключение отметим одну теорему, касающуюся вращения.

IV. Графики G_8 и G_9 уравнений

а) $F\left(\frac{x+y}{\sqrt{2}}, \frac{-x+y}{\sqrt{2}}\right)=0$ и б) $F\left(\frac{x-y}{\sqrt{2}}, \frac{x+y}{\sqrt{2}}\right)=0$

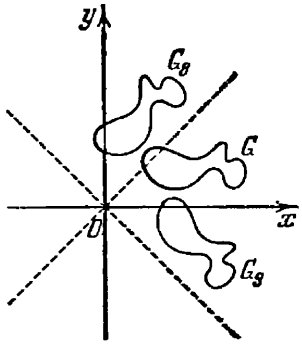


Рис. 10.

получаются из графика G уравнения $F(x, y)=0$ посредством вращения на 45° около начала O в прямом направлении (т. е. от оси Ox к Oy) или обратном (от оси Oy к оси Ox), как показано на рис. 10.

Это следует из общих формул поворота координатной системы на угол φ [1]):

$$\begin{cases} x=x'\cos\varphi-y'\sin\varphi, \\ y=x'\sin\varphi+y'\cos\varphi, \end{cases}$$

если положить $\varphi=\frac{\pi}{4}$.

§ 4. Прямая и обратная функции

В связи с теоремами I—IV предыдущего параграфа поставим ещё вопрос: что делается с графиком уравнения $F(x, y)=0$, если поменять местами буквы x и y?

Оказывается, что имеет место следующая теорема.

V. График G_{10} уравнения

$$F(y, x)=0$$

симметричен графику G уравнения $F(x, y)=0$ относительно биссектрисы $y=x$ первого координатного угла (рис. 11).

В самом деле, если точка $M(x_0, y_0)$ принадлежит графику G, то точка $M'(y_0, x_0)$ принадлежит графику G_{10}; но точка M', очевидно, симметрична точке M относительно упомянутой биссектрисы.

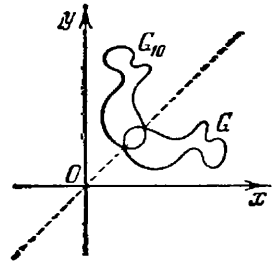


Рис. 11.

В частности, если функция $F(x, y)$ симметрична относительно переменных x и y, т. е. не изменяется при их перестановке

$$F(y, x)\equiv F(x, y),$$

то график уравнения $F(x, y)=0$ имеет прямую $x=y$ осью симметрии.

Примерами могут служить уравнения

$x+y=1$, $x^2+y^2=25$, $xy=1$, $x+y=xy$, $\sqrt{x}+\sqrt{y}=1$ и т. п.

---

[1]) См. Э. э. м., кн. 4 «Геометрия, часть I», статья «Элементы аналитической геометрии».

Уравнение $F(x, y)=0$ позволяет рассматривать y как некоторую (однозначную или многозначную, см. § 1) функцию переменной x. Допустим точно так же, что оно позволяет рассматривать x как некоторую, однозначную или многозначную, функцию переменной y. Этим, вообще говоря, различным (с точки зрения выполняемых операций), функциям нередко дают наименования прямой и обратной.

Ограничиваясь простейшим случаем, предположим, что в пределах некоторой части плоскости (или, может быть, во всей плоскости) каждая прямая, параллельная оси Oy, имеет с графиком уравнения

$$F(x, y)=0 \tag{26}$$

только одну общую точку, и вместе с тем каждая прямая, параллельная оси Ox, также имеет с графиком только одну общую точку. Это значит, что в пределах названной части плоскости уравнение (26) может быть «решено относительно y»

$$y=f(x) \tag{27}$$

и вместе с тем может быть «решено относительно x»

$$x=g(y). \tag{28}$$

При этом мы ограничиваемся здесь предположением, что для определения значения x по данному значению y, или обратно, достаточно элементарных операций, так что функции, обозначенные буквами f и g, — элементарные. Тогда функцию f можно считать прямой, функцию g — обратной (или, если угодно, наоборот). Так как согласно предположению три уравнения (26), (27) и (28) равносильны, то у них один и тот же график — скажем, G.

В уравнении (28) вопреки принятым нами обозначениям роль независимого переменного играет буква y, роль зависимого — буква x. Поменяв местами эти буквы, мы получим такое же уравнение, с точки зрения выполняемых операций:

$$y=g(x), \tag{29}$$

но график его будет уже другой: он согласно изложенному выше симметричен графику G относительно прямой $x=y$.

Мы приходим, таким образом, к важному выводу: *если обозначать независимую переменную одной и той же буквой x, то график прямой функции $y=f(x)$ и график обратной функции $y=g(x)$ взаимно симметричны относительно прямой $x=y$.*

Пример 1. Уравнение $2x-5y+10=0$ (см. § 1) равносильно каждому из уравнений

$$y=\frac{2x+10}{5} \quad \text{или} \quad x=\frac{5y-10}{2}.$$

Если в качестве прямой функции взять

$$f(x)=\frac{2x+10}{5},$$

то обратной функцией будет

$$g(x)=\frac{5x-10}{2}.$$

Пример 2. Уравнению $x^2-y=0$ можно придать вид

$$y=x^2.$$

Решая его относительно x, получаем:

$$x=\begin{cases}+\sqrt{y} & \text{в правой полуплоскости } (x>0),\\ -\sqrt{y} & \text{в левой полуплоскости } (x<0).\end{cases}$$

Итак, если прямая функция есть $f(x)=x^2$, то обратной является $g(x)=\sqrt{x}$ (радикал — в алгебраическом смысле). Или более точно: обратной функцией является $g(x)=+\sqrt{x}$ в верхней полуплоскости $(y>0)$, $g(x)=-\sqrt{x}$ — в нижней $(y<0)$ (здесь радикалы — в арифметическом смысле).

Пример 3. Если $f(x)=\frac{1}{100}(x^3+1)$, то $g(x)=\sqrt[3]{100x-1}$.

Пример 4. По отношению к функции

$$f(x)=\frac{x}{1+x}$$

обратной функцией будет

$$g(x)=\frac{x}{1-x}.$$

Пример 5. Если в качестве прямой функции взять

$$f(x)=\lg\cos x,$$

то обратной функцией будет

$$g(x)=\operatorname{Arccos}10^x.$$

§ 5. Элементарное исследование функций (постановка вопроса и некоторые общие приёмы)

Пусть дано уравнение вида

$$y=f(x). \tag{30}$$

Под словом «исследование» функции $f(x)$ (или её графика G) можно понимать выяснение самых разнообразных свойств этой функции. Однако чаще всего, говоря об «исследовании», имеют в виду найти ответы на следующие вопросы:

А. Обращается ли функция в нуль и в каких именно точках? В каких промежутках она положительна и в каких — отрицательна?

Б. Какова «изменяемость» функции? В каких промежутках функция возрастает и в каких — убывает? В каких точках она имеет наибольшее значение (максимум) и в каких — наименьшее (минимум)?

Термины «возрастает», «убывает», «максимум», «минимум» требуют объяснений.

Говорят, что функция $f(x)$ в данном промежутке возрастает, если, каковы бы ни были два значения x из этого промежутка, например, x' и x'', как только

$$x' < x'', \tag{31}$$

то непременно

$$f(x') < f(x''). \tag{32}$$

(Тогда уже и обратно: из неравенства (32) следует неравенство (31), так как, если бы было $x'' < x'$, то из этого следовало бы, что $f(x'') < f(x')$, а из равенства $x'' = x'$ следовало бы, что $f(x'') = f(x')$.)

Иначе можно сказать: свойство «возрастания» означает, что *с увеличением независимой переменной увеличивается также и функция* (зависимая переменная), *и обратно*; напротив, *с уменьшением независимой переменной уменьшается и функция, и обратно.* Таким образом, переменные x и y изменяются «в одном направлении» [1]).

Говорят, что функция $f(x)$ в данном промежутке убывает, если, каковы бы ни были два числа x' и x'' из этого промежутка, из неравенства

$$x' < x'' \tag{33}$$

непременно следует неравенство

$$f(x') > f(x'') \tag{34}$$

(и тогда, конечно, обратно). Таким образом, в этом случае с увеличением независимой переменной x функция $y = f(x)$ уменьшается, а с уменьшением x — увеличивается; переменные x и y изменяются «в разных направлениях» [2]).

---

[1]) С точки зрения теории неравенств вопрос целесообразно поставить в следующей форме: можно ли утверждать, что из неравенства вида $A < B$ непременно вытекает неравенство $f(A) < f(B)$? Ответ — утвердительный, если только в рассматриваемом промежутке функция $f(x)$ — возрастающая.

Примеры возрастающих функций: $f(x) = mx$ $(m > 0)$, $f(x) = x^3$ для любого промежутка; $f(x) = x^2$, $f(x) = \sqrt{x}$ для промежутка $0 \leqslant x < \infty$.

[2]) Если функция $f(x)$ — убывающая (в рассматриваемом промежутке), то из неравенства $A < B$ следует неравенство $f(A) > f(B)$.

Примеры убывающих функций: $f(x) = mx$ $(m < 0)$ в любом промежутке; $f(x) = \frac{1}{x}$, $\frac{1}{x^s}$ $(s > 0)$ при положительных значениях x.

Если $a<c<b$ и если в промежутке от a до c функция $f(x)$ возрастает, а в промежутке от c до b убывает, то в точке $x=c$ функция $f(x)$ имеет максимум; если, напротив, в промежутке от a до c функция убывает, а в промежутке от c до b возрастает, то в точке $x=c$ она имеет минимум. Сказанное не является определением максимума и минимума: согласно точному, принятому в науке определению функция имеет максимум в точке $x=c$, если можно указать два таких числа a и b, что

1) $a<c<b$,

2) неравенство

$$f(c)>f(x) \tag{35}$$

имеет место для любого x, удовлетворяющего условиям $a<x<b$, $x\neq c$. Аналогично для минимума. В дальнейшем (см. § 46, пример 4) на примере будет обнаружено, что вторая формулировка несколько шире первой.

Геометрический смысл понятий возрастания и убывания, максимума и минимума достаточно очевиден и не требует пояснений.

Примечание 1. Понятия «возрастающая функция» и «убывающая функция» могут быть несколько расширены. Именно, можно предположить, что взамен неравенства (32), как следствие неравенства $x'<x''$, должно выполняться менее стеснительное соотношение

$$f(x')\leqslant f(x''). \tag{32'}$$

Таким образом, предполагается, что с увеличением независимой переменной функция или возрастает или не изменяет значения (но не убывает): такую функцию называют *неубывающей*.

Всякая возрастающая функция есть вместе с тем неубывающая; обратное неверно.

Вот «житейский» пример неубывающей функции: x — время, $f(x)$ — показание электросчётчика. С течением времени при затрате электроэнергии показание счётчика увеличивается; но если в течение какого-нибудь промежутка времени энергия не расходуется, то не изменяется и показание счётчика.

Можно доказать, что всякая элементарная неубывающая функция непременно должна быть возрастающей (если она не сводится к постоянной).

Точно так же функция, удовлетворяющая требованию

$$f(x')\geqslant f(x''), \tag{34'}$$

менее стеснительному, чем требование (34), называется *невозрастающей*.

Примечание 2. Понятия максимума и минимума также допускают некоторое расширение. Говорят, что в точке c функция $f(x)$ имеет максимум в расширенном смысле, если взамен точного неравенства (35) в окрестности этой точки выполняется соотношение

$$f(c)\geqslant f(x). \tag{35'}$$

Аналогично для минимума.

Примечание 3. Изменение величины в одном и том же направлении (в сторону возрастания или в сторону убывания) или состояние неизменности в математике называется *монотонным* изменением. *Монотонными* называются также неубывающие и невозрастающие функции.

Эти термины сообщаются читателю для сведения: мы далеки от мысли о том, что было бы целесообразно вводить их в практику преподавания.

Примечание 4. Ради сокращения речи в математике термины «максимум» (maximum — наибольшее) и «минимум» (minimum — наименьшее) принято объединять в один термин «экстремум» (extremum — крайнее); в преподавание в средней школе этот термин также нет надобности вводить.

По поводу вопросов типа А (см. стр. 34) мы ограничимся здесь общим замечанием о том, что, желая выяснить знак функции $f(x)$, обыкновенно стараются данную формулу представить (подвергая тождественным преобразованиям) в виде произведения таких множителей, чтобы знак каждого из них мог быть легко установлен. Затем остаётся сослаться на элементарные свойства произведений:

1) произведение равно нулю в том и только в том случае, если хотя бы один из множителей равен нулю,

2) если все множители отличны от нуля, то произведение положительно или отрицательно, смотря по тому, будет ли число отрицательных множителей чётным или нечётным.

В иных случаях, когда такого рода разложение на множители не удаётся, пытаются свести вопросы типа А к вопросам типа Б.

Например, установив, что в рассматриваемом промежутке функция возрастает и что в начальной точке промежутка она принимает положительное значение, можно сейчас же заключить, что функция положительна во всём промежутке. Или более общо: к тому же самому заключению можно придти, если удастся установить, что положительным является наименьшее значение функции во всём рассматриваемом промежутке.

Заметим, что точки, в которых рассматриваемая функция принимает значение нуль, очень часто ради краткости и удобства речи называют *нулями* этой функции.

По поводу вопросов типа Б нужно сказать, что они принципиально сложнее: общий метод для их разрешения даётся в дифференциальном исчислении; исследование возрастания и убывания функции составляет одну из основных задач математического анализа. Метод дифференциального исчисления по существу заключается в том, что о возрастании и убывании функции судят по знаку некоторой другой функции [1]), называемой «производной» от функции (см. стр. 307).

Однако существуют и разнообразные, хотя отнюдь не исчерпывающие, элементарные методы исследования изменяемости функции, представляющие к тому же особый интерес с точки зрения школь-

[1]) То-есть вопрос типа Б сводится к вопросу типа А.

ной практики; на некоторых из них здесь уместно остановиться. Следует заметить, что в одних случаях вообще совершенно излишне обращаться к методам анализа, так как вопрос решают наиболее простым и естественным путём элементарными методами; в других же случаях, хотя элементарное решение и возможно, оно оказывается гораздо более сложным, чем решение методами анализа, а подчас носит и вовсе искусственный характер (см., например, конец § 10).

Элементарные методы основываются на том, что «изменяемость» простейших функций исследуется, исходя непосредственно из определения возрастания или убывания. Вместе с тем ряд теорем, почти очевидных, позволяет судить по «изменяемости» одних функций, более простых, об изменяемости других, более сложных.

Условимся ради краткости говорить, что в рассматриваемом промежутке функции $u(x)$ и $v(x)$ изменяются в одном и том же направлении, если обе они — возрастающие или обе — убывающие; напротив, будем говорить, что функции $u(x)$ и $v(x)$ изменяются в противоположных направлениях, если одна из функций — возрастающая, а другая — убывающая.

В таком случае:

1. Функции $u(x)$ и $u(x)+C$ (где C — постоянное число) изменяются в одном и том же направлении.

2 и 2′. Функции $u(x)$ и $Cu(x)$ изменяются в одном и том же или в противоположных направлениях, смотря по тому, является ли C положительной или отрицательной постоянной.

3. Если функции $u(x)$ и $v(x)$ изменяются в одном и том же направлении, то сумма $u(x)+v(x)$ изменяется в том же направлении.

4. Если положительные функции $u(x)$ и $v(x)$ изменяются в одном и том же направлении, то произведение $u(x)v(x)$ изменяется в том же направлении.

4′. Напротив, если отрицательные функции $u(x)$ и $v(x)$ изменяются в одном и том же направлении, то произведение $u(x)v(x)$ изменяется в противоположном направлении.

5 и 5′. В частности, если $u(x)$ — положительная функция, то $u^2(x)$ изменяется в том же направлении, что и $u(x)$; если $u(x)$ — отрицательная функция, то $u^2(x)$ изменяется в направлении противоположном.

6. Функции $u(x)$ и $-u(x)$ изменяются в противоположных направлениях.

7. Если функция $u(x)$ положительна в некотором промежутке, то функции $u(x)$ и $\frac{1}{u(x)}$ в этом промежутке изменяются в противоположных направлениях.

7′. Если функция $u(x)$ отрицательна в некотором промежутке, то $u(x)$ и $\frac{1}{u(x)}$ точно так же изменяются в противоположных направлениях.

Таким образом, относительно всякого промежутка, в котором функция $u(x)$, не обращаясь в нуль, сохраняет один и тот же знак, можно сказать, что функции $u(x)$ и $\frac{1}{u(x)}$ изменяются в противоположных направлениях.

8. Если функция $u(x)$ — положительная, то функции $u(x)$ и $\sqrt{u(x)}$ изменяются в одном и том же направлении (здесь $\sqrt{u}$ — арифметическое значение корня).

9. «Сложная функция» $u(v(x))$ возрастает, если обе функции $u(x)$ и $v(x)$ изменяются в одном и том же направлении; убывает — если $u(x)$ и $v(x)$ изменяются в разных направлениях [1]).

Доказательства всех этих теорем однотипны и крайне элементарны; дело сводится к использованию определения убывания или возрастания функции, с привлечением простейших свойств неравенств. В качестве примера приведём одно из них, именно, относящееся к теореме 3, с допущением возрастания.

Сумма двух возрастающих функций есть функция возрастающая.

Доказательство. Пусть $f(x)=u(x)+v(x)$, причём функции $u(x)$ и $v(x)$ — возрастающие в некотором промежутке. Предположим, далее, что x' и x'' — какие угодно числа из этого промежутка, причём $x'<x''$.

По предположению имеем:

$$u(x')<u(x''),$$
$$v(x')<v(x'').$$

Складывая эти неравенства, получаем

$$u(x')+v(x')<u(x'')+v(x''),$$

т. е.

$$f(x')<f(x'').$$

Теоремами 1—8 в дальнейшем мы будем пользоваться большей частью без соответствующих ссылок.

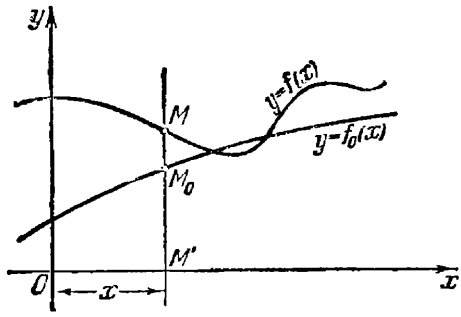


Рис. 12.

В порядке примечания не мешает отметить, что «исследование» функций не всегда ограничивается вопросами, сформулированными в пунктах А и Б; однако оно очень часто может быть сведено к ответу на эти вопросы.

Приведём пример. Пусть, кроме исследуемого графика G функции $y=f(x)$, имеется уже хорошо известный нам график G_0 функции $y=f_0(x)$ (рис. 12). Нас могли бы интересовать такие вопросы:

---

[1]) Точнее: нужно рассматривать изменяемость функции $u(x)$ в том промежутке, какому принадлежат значения функции $v(x)$ в процессе изменения переменной x.

А′. Пересекается ли график G с графиком G_0 и в каких именно точках? В каких промежутках график G лежит выше графика G_0 и в каких — ниже?

Б′. Как изменяется (с увеличением x) направленный вертикальный отрезок M_0M между точкой пересечения с графиком G_0 и точкой пересечения с графиком G? В каких промежутках он возрастает? В каких убывает? В каких точках принимает наибольшее и в каких — наименьшее значение?

Легко понять, что вопросы А′ и Б′ сводятся к вопросам А и Б, если только вместо самой функции $f(x)$ ввести разность $f(x) - f_0(x)$.

При исследовании различных свойств данной функции формулу, с помощью которой она задана, часто приходится подвергать тем или иным тождественным преобразованиям. *Выбор преобразования в каждом данном случае зависит от того свойства функции, которое имеется в виду установить.*

ГЛАВА II

ОБЗОР ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ И ИХ ГРАФИКОВ

§ 6. Классификация рациональных функций

Рациональными называются такие функции, значения которых можно найти, не совершая над заданным значением независимой переменной x и числами, которые предполагаются постоянными, никаких операций, кроме четырёх действий арифметики (сложение, вычитание, умножение, деление). Из числа рациональных функций *целыми*[1]) называются такие, значения которых можно получить из переменной x и постоянных чисел посредством не более чем трёх операций: сложения, вычитания и умножения. Упоминание о вычитании не обязательно, так как вычесть число или буквенное выражение — значит прибавить его, предварительно умножив на -1. Те рациональные функции, которые не являются целыми, носят название *дробных*. Примеры целых рациональных функций:

$$2x-3,\quad (x-1)(x+4),\quad x^3,\quad \frac{x}{5}\left(=x\cdot\frac{1}{5}\right),$$

$$\sqrt{x^4+4x^2+4}\quad (=x^2+2)^{2}),\quad \frac{x^2-1}{x-1}\ (=x+1),$$

$$\cos(2\arccos x)\ (=2x^2-1)^{3}).$$

Примеры дробных рациональных функций:

$$\frac{x+1}{x^2-2x+5},\quad x+\frac{1}{x},\quad x^{-3}\left(=\frac{1}{x^3}\right),\quad \operatorname{tg} 3\operatorname{arctg} x\left(=\frac{3x-x^3}{1-3x^2}\right).$$

Методом полной математической индукции легко установить:

1) Всякая целая рациональная функция $P(x)$ может быть представлена в виде многочлена, расположенного по убывающим степеням x:

$$P(x)=ax^n+bx^{n-1}+\ \ldots\ +kx+l\qquad (a\neq 0;\ n\text{ — целое, }\geqslant 0).$$

Число n называется *степенью* многочлена $P(x)$.

---

[1]) Вместо «целая рациональная функция» принято говорить короче «рациональный многочлен», или «полином».

[2]) Имеется в виду арифметическое значение корня.

[3]) Ограничение $|x|\leqslant 1$ (которое кажется необходимым) будет снято в статье «Элементарные функции комплексного переменного», см. стр. 511.

2) Всякая дробная рациональная функция $R(x)$ может быть представлена в виде отношения двух многочленов, расположенных по убывающим степеням x (без общих множителей):

$$R(x) = \frac{P(x)}{Q(x)}.$$

Под *степенью дробной* рациональной функции понимают наибольшую из степеней числителя $P(x)$ и знаменателя $Q(x)$ [1]).

При доказательстве предложения 1) индукция производится по числу операций. Обозначим это число через N. Заметим предварительно, что а) сумма, б) разность, в) произведение двух многочленов, расположенных по убывающим степеням x, могут быть также представлены в виде многочлена, расположенного по убывающим степеням x. Допустим теперь, что предложение 1) справедливо для всякой целой рациональной функции, «составленной» из переменной x и постоянных посредством N операций, и докажем, что оно будет справедливо также для целой рациональной функции, «составленной» таким же образом посредством $N+1$ операций. Последняя $(N+1)$-я операция, указываемая выражением, с помощью которого задана наша функция, не может быть ничем иным, как только сложением, вычитанием или умножением; но так как каждая из двух целых рациональных функций, над которыми нужно произвести названную операцию, «составлена» не более как из N операций, то каждая из них может быть представлена в виде многочлена, расположенного по убывающим степеням x. Значит, на основании предварительного замечания данная функция также может быть представлена в виде многочлена, расположенного по степеням x.

Остаётся заметить, что для случая $N=1$ утверждение, очевидно, справедливо.

Предложение 2) доказывается также посредством индукции, но уже по числу делений. Предварительно нужно заметить, что а) сумма, б) разность, в) произведение, г) отношение двух дробей, у которых числитель и знаменатель — многочлены, также могут быть представлены в виде дроби того же вида.

§ 7. Целые положительные степени

Рассмотрим графики функций

$$x,\ x^2,\ x^3,\ \ldots,\ x^n,\ \ldots \tag{1}$$

График G_1 функции

$$y = x \tag{2}$$

— совокупность точек, у которых абсцисса равна ординате, представляет собой биссектрису первого и третьего координатных углов,

[1]) Следует не упускать из виду различие между степенью рациональной функции и порядком алгебраической кривой — графика этой функции (см. ниже § 17). С геометрической точки зрения порядок кривой есть максимальное число точек её пересечения с произвольной прямой, а степень функции — максимальное число точек пересечения её графика с произвольной горизонтальной прямой (включая и «мнимые» точки пересечения). Для целых функций степень и порядок графика совпадают, для дробных они могут различаться. Например, дробная функция $y = \frac{1}{x}$ — первой степени, а её график — гипербола — кривая второго порядка.

т. е. *прямую линию*, проходящую через начало O и образующую угол в 45° с осью Ox (рис. 13, *а*).

График G_2 функции

$$y = x^2 \tag{3}$$

(рис. 13, *б*) носит название *параболы*. (В расширенном смысле параболами называют иногда графики произвольных рациональных мно-

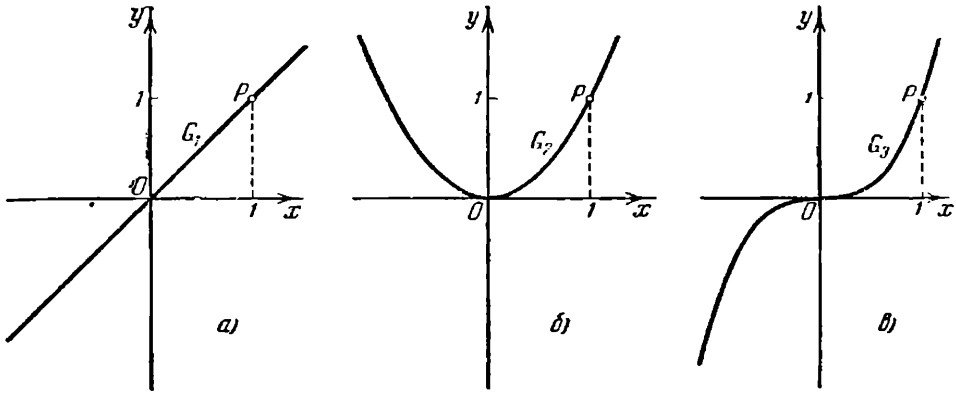


Рис. 13.

гочленов.) Он симметричен относительно оси Oy, так как функция x^2 — чётная. Он весь лежит в верхней полуплоскости (так как $x^2 \geqslant 0$), кроме одной точки — начала O, *вершины* параболы, в которой кривая пересекается со своей осью симметрии. Функция x^2 — очевидно, возрастающая; значит, при $x > 0$ функция x^2 — также возрастающая. Напротив, при $x < 0$ она — убывающая по свойству симметрии (или вследствие предложения 4′ на стр. 38). Чтобы выяснить взаимное расположение графиков (2) и (3), рассмотрим разность $x^2 - x$. Разлагая её на множители

$$x^2 - x = x\,(x - 1),$$

видим, что графики G_2 и G_1 имеют две общие точки: начало O (0, 0) и точку P (1, 1). При этом, если $x > 1$, то $x^2 - x > 0$, так что G_2 лежит выше G_1; если $0 < x < 1$, то $x^2 - x < 0$ и G_2 ниже G_1; если $x < 0$, то ясно, что G_2 выше G_1. При очень малых значениях $|x|$ отношение

$$\frac{y}{x} = x$$

очень мало́ по абсолютной величине, и потому около вершины кривая тесно примыкает к горизонтальной оси («касается» её, см. стр. 309). При очень больших значениях $|x|$ это отношение, напротив, очень велико, что говорит о быстром возрастании x^2.

График G_3 функции

$$y = x^3 \tag{4}$$

(«кубическая парабола») имеет центр симметрии O (рис. 13, *в*), так как функция x^3 — нечётная. Подобно G_1, он расположен в первом и третьем квадрантах. С графиком G_2 он имеет общие точки O и P, причём лежит ещё ниже его в промежутке $0 < x < 1$ и ещё выше

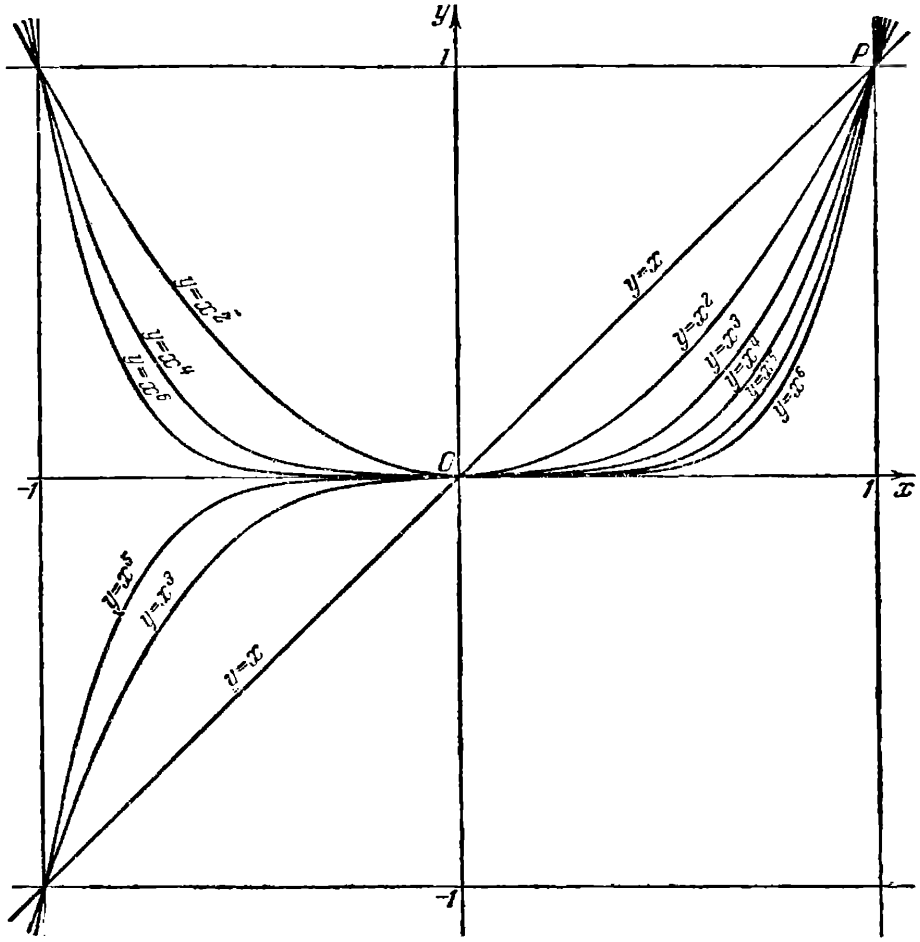


Рис. 14.

его в промежутке $1 < x < \infty$. Около начала O он примыкает к оси Ox ещё теснее, чем G_2 (в точке O имеется «касание второго порядка», причём касательная пересекает кривую); при неограниченном возрастании x функция x^3 растёт еще быстрее, чем x^2.

Обозначая вообще через G_n график функции

$$y = x^n$$

(при любом целом положительном n), можно установить, что он проходит через точки O и P; при n чётном он имеет ось симметрии Oy

и располагается в первом и втором квадрантах, а при n нечётном имеет центр симметрии O и располагается в первом и третьем квадрантах. В промежутке $0 < x < 1$ значения функций (1) при данном значении x образуют убывающую геометрическую прогрессию: из последовательности графиков

$$G_1, G_2, G_3, \ldots, G_n, \ldots \tag{5}$$

каждый следующий лежит ниже предыдущего, всё теснее примыкая к горизонтальной оси; в промежутке $1 < x < \infty$ эти же значения образуют возрастающую геометрическую прогрессию, и из графиков (1) каждый следующий лежит выше предыдущего.

На рис. 14 изображены графики функций x, x^2, x^3, x^4, x^5, x^6 из совокупности (1).

§ 8. Многочлены первой степени (линейные функции)

Линейная функция имеет общий вид

$$y = ax + b \qquad (a \neq 0). \tag{6}$$

Из уравнения видно, что её график может быть получен из графика прямой $y = x$ посредством растяжения в $|a|$ раз по направлению оси Oy, с последующим (в случае $a < 0$) симметричным отражением относительно оси Oy и затем — перенесением параллельно оси Oy на отрезок, равный b.

В самом деле, рассмотрим последовательно графики уравнений

$$1)\ y = x,$$
$$2)\ y = |a|\, x,$$
$$3)\ y = ax,$$
$$4)\ y = ax + b.$$

График 1) нам известен: это — биссектриса 1-го и 3-го координатных углов. Уравнению 2) можно придать вид $\frac{y}{|a|} = x$, откуда ясно (см. § 3, III. б), что его график получается из графика 1) посредством растяжения в $|a|$ раз по направлению оси Oy. График 3) в случае $a > 0$ ничем не отличается от графика 2); в случае $a < 0$ уравнению 3) можно придать вид $y = -|a|\, x$ или $(-y) = |a|\, x$, откуда ясно, что график 3) получается в этом случае из графика 2) посредством симметричного отражения относительно оси Ox. Наконец, легко понять (§ 3, II. б), что график 4) получается из графика 3) посредством перенесения параллельно оси Oy на отрезок, равный b [1]).

[1]) В дальнейшем рассуждения, подобные предыдущему, будут опускаться, и вместо них будут приведены лишь ссылки на пункты § 3.

Запись

$$y = a\left(x + \frac{b}{a}\right)$$

показывает, что график уравнения (6) может быть также получен посредством перенесения графика $y = x$ параллельно оси Ox на отрезок, равный $\left(-\frac{b}{a}\right)$, и последующего растяжения в $|a|$ раз по направлению оси Oy, с симметричным отражением (в случае $a < 0$) относительно оси Ox.

Так или иначе, ясно, что график линейной функции (6) — наклонная (т. е. не параллельная ни одной из координатных осей) прямая линия.

Коэффициент a в уравнении (6) называется *наклоном* (или *подъёмом*) прямой: его называют также *угловым коэффициентом*. Он представляет собой тангенс угла, который прямая образует с положительным направлением оси Ox: действительно, если $a = 1$, это очевидно, так как в этом случае упомянутый угол составляет 45°; с другой стороны, при растяжении в $|a|$ раз по направлению оси Oy тангенс увеличивается во столько же раз, а при отражении относительно оси Ox меняется знак как угла, так и его тангенса.

Свободный член b представляет собой значение функции при $x = 0$, т. е. «отрезок, который прямая отсекает на оси Oy», или, точнее, ординату точки пересечения прямой с этой осью.

Что касается точки пересечения с осью Ox, то её абсцисса равна $\left(-\frac{b}{a}\right)$.

Рассматриваемая функция, как легко понять, — возрастающая, если $a > 0$, и убывающая, если $a < 0$.

В случае $a = 0$ линейная функция сводится к постоянной («многочлен нулевой степени»); её график — прямая, параллельная оси Ox.

§ 9. Многочлены (трёхчлены) второй степени

Трёхчлен второй степени имеет общий вид:

$$y = ax^2 + bx + c \qquad (a \neq 0). \tag{7}$$

Выполним в правой части следующие преобразования: вынесем a за скобки, затем «выделим точный квадрат». Запись

$$y = a\left[\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2}\right] \tag{8}$$

показывает, что график уравнения (7) получается из параболы $y = x^2$ посредством последовательного выполнения следующих преобразований:

1) перенесения параллельно оси Ox на отрезок $\left(-\frac{b}{2a}\right)$;

2) перенесения параллельно оси Oy на отрезок $\frac{4ac-b^2}{4a^2}$;

3) растяжения в $|a|$ раз по направлению оси Oy с отражением (в случае $a<0$) относительно оси Ox.

Уравнению (7) можно также придать вид

$$y=a\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a}, \tag{9}$$

и отсюда следует, что возможен иной план преобразований:

1) перенесение параллельно оси Ox на отрезок $\left(-\frac{b}{2a}\right)$,

2) растяжение в $|a|$ раз по направлению оси Oy с отражением (в случае $a<0$) относительно оси Ox,

3) перенесение параллельно оси Oy на отрезок $\frac{4ac-b^2}{4a}$.

Таким образом, график трёхчлена второй степени, заданного общим уравнением (7), есть парабола, конгруэнтная параболе

$$y=|a|x^2,$$

получающейся из основной параболы $y=x^2$ посредством растяжения в $|a|$ раз по направлению оси Oy.

Независимо от описанных геометрических преобразований, из записи (8) или (9) видно (см. § 5), что в случае $a>0$ функция ax^2+bx+c при $x<-\frac{b}{2a}$ убывает, при $x>-\frac{b}{2a}$ возрастает, в точке $x=-\frac{b}{2a}$ имеет минимум, значение которого равно $\frac{4ac-b^2}{4a}$; в случае $a<0$, напротив, при $x<-\frac{b}{2a}$ она возрастает, при $x>-\frac{b}{2a}$ убывает, в точке $x=-\frac{b}{2a}$ имеет максимум, значение которого равно $\frac{4ac-b^2}{4a}$.

При условии, что b отлично от нуля, перенесение параболы совершается влево, если знаки b и a одинаковы; вправо — если эти знаки различны. При b, равном нулю, горизонтальное перенесение отсутствует.

Вертикальное перенесение отсутствует в том случае, если выражение

$$D=4ac-b^2$$

(*дискриминант* трёхчлена) оказывается равным нулю; тогда, как показывает формула (8) или (9), функция ax^2+bx+c сохраняет при всех значениях x, кроме $x=-\frac{b}{2a}$, один и тот же знак, именно — тот же, что и коэффициент a, и лишь при $x=-\frac{b}{2a}$ обращается в нуль. Парабола касается оси Ox.

Если D положительно, то сумма в квадратных скобках в формуле (8) при всех значениях x положительна, и значит, наша функция сохраняет тот же знак, что имеет a, при всех значениях x без исключения. Парабола не пересекается с осью Ox.

Если D отрицательно, то выражение в квадратных скобках можно рассматривать как «разность квадратов», и тогда, полагая

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \quad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

можно представить трёхчлен в виде произведения

$$ax^2 + bx + c = a(x - x_1)(x - x_2) \qquad (x_1 < x_2).$$

Отсюда видно, что трёхчлен обращается в нуль в точках $x = x_1$ и $x = x_2$. Значит, парабола в этих двух точках пересекает ось Ox. Что касается знака трёхчлена, то он не при всех значениях x один и тот же: в случае $a > 0$ знак положителен при $x > x_2$ и $x < x_1$, и отрицателен при $x_1 < x < x_2$; при $a < 0$ — наоборот. Всё это легко истолковывается геометрически.

§ 10. Многочлены третьей степени

Общий вид многочлена третьей степени:

$$y = ax^3 + bx^2 + cx + d \qquad (a \neq 0). \tag{10}$$

Вынесем a за скобки и «выделим точный куб»

$$\begin{aligned} y &= a\left[\left(x + \frac{b}{3a}\right)^3 + \left(\frac{c}{a} - \frac{1}{3}\frac{b^2}{a^2}\right)x + \left(\frac{d}{a} - \frac{1}{27}\frac{b^3}{a^3}\right)\right] = \\ &= a\left[\left(x + \frac{b}{3a}\right)^3 + \left(\frac{c}{a} - \frac{1}{3}\frac{b^2}{a^2}\right)\left(x + \frac{b}{3a}\right) + \left(\frac{d}{a} - \frac{1}{3}\frac{bc}{a^2} + \frac{2}{27}\frac{b^3}{a^3}\right)\right] = \\ &= a\left[\left(x + \frac{b}{3a}\right)^3 + D\left(x + \frac{b}{3a}\right) + D'\right], \end{aligned} \tag{11}$$

где ради краткости положено

$$D = \frac{c}{a} - \frac{1}{3}\frac{b^2}{a^2},$$

$$D' = \frac{d}{a} - \frac{1}{3}\frac{bc}{a^2} + \frac{2}{27}\frac{b^3}{a^3}.$$

Далее, в случае $D \neq 0$ можно преобразовать данное выражение ещё следующим образом:

$$y = a|D|^{3/2}\left\{\left(\frac{x + \frac{b}{3a}}{\sqrt{|D|}}\right)^3 \pm \left(\frac{x + \frac{b}{3a}}{\sqrt{|D|}}\right) + \frac{D'}{|D|^{3/2}}\right\}, \tag{12}$$

причём следует перед вторым членом взять знак $+$ или $-$, смотря по тому, будет ли $D > 0$ или $D < 0$.

Если же $D=0$, то формула (11) принимает вид

$$y=a\left\{\left(x+\frac{b}{3a}\right)^3+D'\right\}. \tag{13}$$

Глядя на формулы (12) и (13), можно сделать заключение, что посредством элементарных преобразований (см. § 3, I—III) график данного многочлена третьей степени общего вида может быть получен из графиков следующих простейших многочленов:

$$(\alpha)\ y=x^3+x,$$
$$(\beta)\ y=x^3-x,$$
$$(\gamma)\ y=x^3.$$

Именно, в случае $D\neq 0$ данный график получается из графиков (α) или (β) (смотря по тому, будет ли $D>0$ или $D<0$) посредством следующих преобразований:

1) растяжения в $\sqrt{|D|}$ раз по направлению оси Ox,

2) перенесения параллельно оси Ox на отрезок $\left(-\frac{b}{3a}\right)$,

3) перенесения параллельно оси Oy на отрезок $\frac{D'}{|D|^{3/2}}$,

4) растяжения в $|a|\cdot|D|^{3/2}$ раз по направлению оси Oy, с отражением (в случае $a<0$) относительно оси Ox.

В случае же $D=0$ рассматриваемый график получается из графика (γ) посредством

1) перенесения параллельно оси Ox на отрезок $\left(-\frac{b}{3a}\right)$,

2) перенесения параллельно оси Oy на отрезок D',

3) растяжения в $|a|$ раз по направлению оси Oy, с отражением (в случае $a<0$) относительно оси Ox.

Итак, чтобы отдать себе отчёт в расположении графиков функций вида (10) и в случае надобности провести их элементарное исследование, достаточно рассмотреть кривые (α) и (β); что касается кривой (γ), то о ней уже было сказано раньше (см. § 7).

Обе функции (α) и (β) — нечётные, так что графики их (рис. 15) симметричны относительно центра O. При $x>0$ функция (α) — возрастающая (см. § 5), и её график, очевидно, лежит выше прямой $y=x$, тесно к ней примыкая около начала O (касаясь её в этой точке). Поведение графика (β) сложнее: разложение на множители

$$x^3-x=x(x-1)(x+1)$$

показывает, что имеются пересечения с осью Ox в точках $x=0$, $x=1$ и $x=-1$. Ограничимся рассмотрением положительных значений x. Мы видим, что $x^3-x>0$ при $x>1$, но $x^3-x<0$ при $0<x<1$.

Вместе с тем непосредственно ясно, что график лежит выше прямой $y=-x$ и касается её в начале координат. Из преобразования

$$x^3-x=-\frac{2}{3\sqrt{3}}+\left(x-\frac{1}{\sqrt{3}}\right)^2\left(x+\frac{2}{\sqrt{3}}\right). \qquad (14)$$

можно заключить, что при $x=\frac{1}{\sqrt{3}}$ функция имеет минимум, равный $\left(-\frac{2}{3\sqrt{3}}\right)$; отсюда же видно, что она возрастает при $x>\frac{1}{\sqrt{3}}$ (см. § 5).

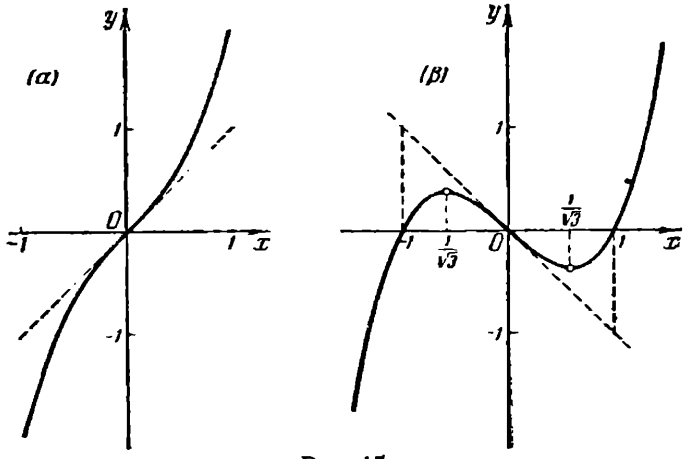


Рис. 15.

Доказать элементарно, что функция $f(x)=x^3-x$ убывает в промежутке $0<x<\frac{1}{\sqrt{3}}$, можно проще всего, основываясь непосредственно на определении. Считая, что величины

$$x'=x$$

и

$$x''=x+h$$

удовлетворяют неравенствам

$$0\leqslant x'<x''\leqslant\frac{1}{\sqrt{3}},$$

так что

$$0<h\leqslant\frac{1}{\sqrt{3}},$$

$$0\leqslant x\leqslant\frac{1}{\sqrt{3}}-h,$$

мы получаем

$$f(x'')-f(x')=f(x+h)-f(x)=$$
$$=[(x+h)^3-(x+h)]-(x^3-x)=-h\{1-(3x^2+3hx+h^2)\},$$

и так как

$$3x^2+3hx+h^2\leqslant 3\left(\frac{1}{\sqrt{3}}-h\right)^2+3h\left(\frac{1}{\sqrt{3}}-h\right)+h^2=$$
$$=1-h\sqrt{3}+h^2=1-h(\sqrt{3}-h)<1,$$

то выражение в фигурных скобках положительно и, следовательно, разность $f(x'')-f(x')$ имеет знак, противоположный знаку $h=x''-x'$ [1]).

§ 11. Биквадратные многочлены

Из многочленов четвёртой степени мы остановимся только на *биквадратных*, т. е. на многочленах вида

$$y=ax^4+bx^2+c \quad (a\neq 0). \tag{15}$$

Такие многочлены — чётные функции: их графики симметричны относительно оси Oy.

Как показывает преобразование

$$y=\frac{b^2}{a}\left\{\left(\frac{x}{\sqrt{\left|\frac{b}{a}\right|}}\right)^4\pm\left(\frac{x}{\sqrt{\left|\frac{b}{a}\right|}}\right)^2\right\}+c, \tag{16}$$

в случае $b\neq 0$ график многочлена (15) получается из кривых

$$(\alpha)\ y=x^4+x^2,$$
$$(\beta)\ y=x^4-x^2$$

посредством одних лишь растяжений по направлению обеих осей и перенесения параллельно оси Oy (с отражением относительно оси Ox или без отражения); в случае $b=0$ то же получается, очевидно, из кривой

$$(\gamma)\ y=x^4.$$

Так как функции x^2 и x^4 положительные и возрастающие при $x>0$, причём графики их в начале O касаются оси Ox, то и функция (α) — их сумма — обладает теми же свойствами (рис. 16).

Что касается функции (β), то она, как видно из разложения

$$x^4-x^2=x^2(x-1)(x+1),$$

обращается в нуль в точке $x=0$ (график касается оси Ox) и в точках $x=\pm 1$ (см. рис. 16).

Далее, функцию (α) можно представить в виде $y=\varphi(x^2)$, где положено $\varphi(u)=u^2-u$. При возрастании u от 0 до $\frac{1}{2}$ функция

[1]) В предыдущем изложении не указано, откуда взялось преобразование (14). Это свидетельствует об искусственном характере употребляемого нами элементарного метода. Полезно обратить внимание на то обстоятельство, что дифференциальное исчисление даёт единый и вполне естественный метод для нахождения максимумов и минимумов, а также чрезвычайно облегчает исследование изменяемости функций.

$\varphi(u)$ убывает от 0 до $-\frac{1}{4}$, а при возрастании u от $\frac{1}{2}$ до бесконечности она возрастает от $-\frac{1}{4}$ до бесконечности[1]); значит (см. § 5)

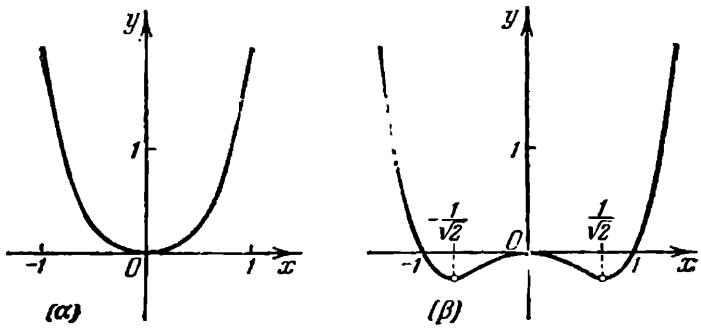


Рис. 16.

при возрастании $x\,(=\sqrt{u})$ от 0 до $\frac{1}{\sqrt{2}}$ функция $y=\varphi(x^2)$ убывает от 0 до $-\frac{1}{4}$, а при возрастании x от $\frac{1}{\sqrt{2}}$ до бесконечности — возрастает от $-\frac{1}{4}$ до бесконечности.

§ 12. Многочлены высших степеней

Главная трудность при исследовании знака многочленов вида

$$y=ax^n+bx^{n-1}+\ldots+kx+l \qquad (a\neq 0;\ n\geqslant 4) \qquad (17)$$

заключается в их разложении на множители, что в свою очередь связано с нахождением всех (по крайней мере действительных) корней алгебраического уравнения

$$ax^n+bx^{n-1}+\ldots+kx+l=0. \qquad (18)$$

Допустим (не ограничивая существенно общности), что старший коэффициент a равен 1.

Из алгебры известно, что после объединения попарно-сопряжённых множителей рассматриваемый многочлен может быть представлен в виде произведения

$$y=(x-\alpha)^{\lambda}(x-\beta)^{\mu}(x-\gamma)^{\nu}\ldots(x^2+px+q)^{\rho}\ldots, \qquad (19)$$

где α, β, ... — действительные, различные между собой, корни уравнения, λ, μ, ... — их кратности (целые положительные числа), и трёхчлены второй степени вида $x^2+px+q\ldots$ имеют сопряжённые мнимые корни.

---

[1]) В самом деле (см. § 9),

$$u^2-u=\left(u-\frac{1}{2}\right)^2-\frac{1}{4}.$$

Так как трёхчлены вида x^2+px+q постоянно сохраняют положительный знак, то знак многочлена y зависит исключительно от знаков множителей вида $(x-\alpha)^\lambda$.

Предположим для определённости, что корни $\alpha, \beta, \ldots$ расположены в порядке убывания: $\alpha>\beta>\gamma>\ldots$; представим себе теперь, что независимая переменная x, постепенно убывая, проходит через значения $x=\alpha$, $x=\beta, \ldots$ Пока x больше, чем α, все множители нашего произведения положительны; значит, положительным остаётся и сам многочлен. Но когда x прошло через значение α, разность $x-\alpha$ становится отрицательной; сохраняет ли при этом множитель $(x-\alpha)^\lambda$ прежний знак или меняет его — зависит от того, является ли кратность λ чётным или нечётным числом. Так продолжается и дальше, при прохождении x через значение β и т. д.

Из этих соображений вытекает такой результат, касающийся графика нашего многочлена: график располагается выше оси Ox при $x>\alpha$; он расположен выше или ниже оси Ox при $\beta<x<\alpha$, смотря по тому, является ли кратность λ чётной или нечётной. При $\gamma<x<\beta$ это уже зависит от чётности $\lambda+\mu$ и т. д. Сумма степеней всех множителей в формуле (19) $\lambda+\mu+\nu+\ldots+2\rho+\ldots$ равна степени n данного многочлена. Если n — чётное, то левее наименьшего из корней (как и правее наибольшего) график лежит выше оси Ox; если n — нечётное, то левее наименьшего из корней он лежит ниже оси Ox. Это справедливо при любом положительном коэффициенте a; при отрицательных a всё происходит наоборот.

Отметим, что в точках, являющихся простыми (т. е. первой кратности) корнями данного многочлена, график пересекает ось Ox без касания; в точках, являющихся корнями чётной кратности, имеет место касание без пересечения с осью; а в точках, являющихся корнями нечётной кратности, начиная с третьей, имеет место касание с пересечением (как и для функции $y=x^n$). Строгое доказательство читатель легко мог бы провести, пользуясь методами дифференциального исчисления.

Возрастание и убывание многочлена общего вида поддаётся элементарному исследованию лишь в частных случаях. Преобразование

$$y=ax^n\left(1+\frac{b}{x}+\ldots+\frac{k}{x^{n-1}}+\frac{l}{x^n}\right) \qquad (20)$$

показывает, что при неограниченном возрастании $|x|$ абсолютная величина функции y также возрастает неограниченно, и тем быстрее, чем больше степень n.

При исследовании многочленов возможны упрощения в том случае, если данный многочлен представляет собой чётную или нечётную функцию. Легко понять, что многочлен есть чётная функция переменной x в том (и только в том) случае, если он содержит лишь чётные степени x, и нечётная функция переменной x в том (и только в том) случае, если содержит лишь нечётные степени x,

§ 13. Целые отрицательные степени

Рассмотрим теперь графики G_1, G_2, G_3, ..., G_n, ... простейших дробных функций

$$\frac{1}{x}, \frac{1}{x^2}, \frac{1}{x^3}, \ldots, \frac{1}{x^n}, \ldots$$

Начнём с функции

$$y=\frac{1}{x}, \tag{21}$$

возникающей при решении уравнения второго порядка

$$xy=1. \tag{22}$$

Её график (рис. 17, *а*) можно построить, исходя из графика прямой $y=x$, в соответствии с указаниями 6, § 2. Имеется центральная

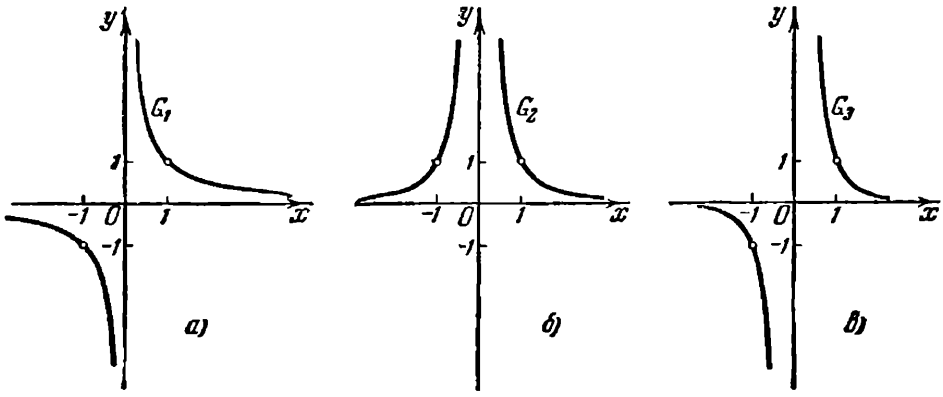


Рис. 17.

симметрия относительно начала координат O; имеется также симметрия относительно прямой $y=x$ (см. § 4); точка $x=0$ — особенная в том смысле, что ей не соответствует никакое значение y: на оси Oy нет ни одной точки графика. Эта точка называется «точкой разрыва» (см. гл. III, § 42); в ней функция $y=\frac{1}{x}$ «теряет смысл». Пусть $x>0$. Так как функция $y=x$ — возрастающая, то функция $y=\frac{1}{x}$, являющаяся величиной, ей обратной, — убывающая. График проходит через точку (1, 1); с неограниченным увеличением x обратная величина $\frac{1}{x}$ становится неограниченно малой — график приближается к оси Ox (как говорят, «асимптотически»); с неограниченным уменьшением x обратная величина $\frac{1}{x}$ неограниченно увеличивается — график, приближаясь к оси Oy, поднимается вверх, «уходит в бесконечность». Можно сказать иначе: уходя в бес-

конечность, график приближается к оси «асимптотически». Вследствие симметрии кривая в целом состоит из двух «ветвей»: одна из них лежит в первом, другая в третьем квадранте (см. рис. 17, *а*). Название кривой — *гипербола*. Оси Ox и Oy — её асимптоты [1]).

Перейдём к функции

$$y=\frac{1}{x^2}, \tag{23}$$

которая может быть рассматриваема как обратная величина по отношению к функции $y=x^2$. Симметрии относительно биссектрисы $y=x$ в данном случае уже нет; при $x>1$ мы имеем $x^2>x$ и $\frac{1}{x^2}<\frac{1}{x}$, так что график G_2 лежит ниже графика G_1 (теснее примыкает к оси Ox); при $0<x<1$ мы имеем, напротив, $x^2<x$ и $\frac{1}{x^2}>\frac{1}{x}$, так что G_2 лежит выше G_1 (поднимается кверху быстрее). Вместо центральной симметрии имеется осевая симметрия относительно оси Oy: таким образом, график G_2 состоит из двух «ветвей», расположенных соответственно в первом и втором квадрантах (рис. 17, *б*).

Рассматривая следующие функции $y=\frac{1}{x^n}$ и их графики G_n $\left(n=3,\ 4,\ \ldots\ ;\right.$ график функции $y=\frac{1}{x^3}$ изображён на рис. 17, *в*$\left.\right)$, мы видим, что все эти функции, как и две уже рассмотренные, «теряют смысл» при $x=0$, являются чётными или нечётными, смотря по чётности n, и при $x>0$ обладают общим свойством — убывать неограниченно при неограниченном возрастании x и возрастать неограниченно при его неограниченном убывании. Так как

$$x<x^2<x^3<\ldots \quad \text{при } x>1$$

и

$$x>x^2>x^3>\ldots \quad \text{при } 0<x<1,$$

то отсюда следует, что

$$\frac{1}{x}>\frac{1}{x^2}>\frac{1}{x^3}>\ldots \quad \text{при } x>1$$

и

$$\frac{1}{x}<\frac{1}{x^2}<\frac{1}{x^3}<\ldots \quad \text{при } 0<x<1.$$

Геометрически всё это означает: каково бы ни было n, в первом квадранте кривая G_{n+1} при $x>1$ лежит ниже кривой G_n, а при

---

[1]) См. Э. э. м., кн. 4, Геометрия, ч. 1, статья «Элементы аналитической геометрии».

$0 < x < 1$ — выше G_n. Все они «асимптотически» приближаются к осям Ox и Oy (рис. 18).

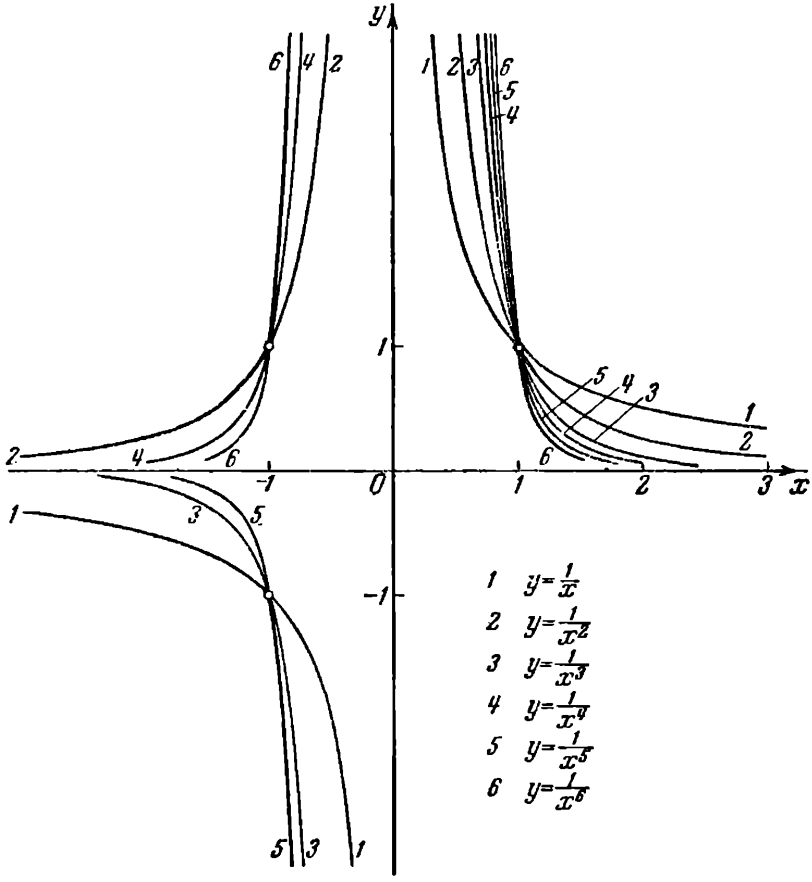


Рис. 18.

Каждая из кривых G_n имеет ещё вторую ветвь, симметричную первой (относительно Oy или O) и лежащую, смотря по чётности n, во втором или третьем квадрантах (см. рис. 18).

Кривые G_n ($n = 3, 4, \ldots$) иногда называются *гиперболами высших порядков.*

§ 14. Дробные линейные функции

Предположим, что данная дробь

$$y = \frac{ax + b}{cx + d} \tag{24}$$

несократима ($ad - bc \neq 0$) и не сводится к целой линейной функции ($c \neq 0$). Тогда, выделяя целую часть и вынося затем за

скобки коэффициент при x в знаменателе, ей можно придать вид

$$y=\frac{a}{c}-\frac{ad-bc}{c^2\left(x+\frac{d}{c}\right)}. \tag{25}$$

Эта запись свидетельствует о том, что график функции $\frac{ax+b}{cx+d}$ можно получить из графика $\frac{1}{x}$ посредством следующих элементарных преобразований:

1) растяжения в $\left|\frac{ad-bc}{c^2}\right|$ раз в направлении оси Oy, с отражением (в случае $ad-bc>0$) относительно оси Ox,

2) перенесения параллельно оси Ox на отрезок, равный $\left(-\frac{d}{c}\right)$,

3) перенесения параллельно оси Oy на отрезок, равный $\frac{a}{c}$.

Таким образом, уравнение (24) также представляет гиперболу (обыкновенную, второго порядка).

Функция $\frac{ax+b}{cx+d}$ «теряет смысл» (не имеет никакого значения, перестаёт «существовать») в точке $x=-\frac{d}{c}$. В промежутках $\left(-\infty, -\frac{d}{c}\right)$ и $\left(-\frac{d}{c}, +\infty\right)$ она сплошь возрастает (если $ad-bc>0$) или сплошь убывает (если $ad-bc<0$)[1]). При неограниченном увеличении $|x|$ значения функции $\frac{ax+b}{cx+d}$ неограниченно приближаются к $\frac{a}{c}$: это видно из преобразования (25) или, проще, из преобразования

$$\frac{ax+b}{cx+d}=\frac{a+\frac{b}{x}}{c+\frac{d}{x}}.$$

Точка $\left(-\frac{d}{c}, \frac{a}{c}\right)$ (не принадлежащая кривой!) представляет собой её *центр* (центр симметрии); прямые $x=-\frac{d}{c}$ и $y=\frac{a}{c}$ — её асимптоты. Итак, для нахождения асимптот кривой (24) достаточно выполнить преобразование (25).

---

[1]) Мы будем пользоваться сокращённым оборотом речи «функция возрастает (или убывает) всюду, кроме точки разрыва».

§ 15. Дробные функции второй степени

Исследование изменяемости несократимых дробей вида

$$y=\frac{ax^2+bx+c}{dx^2+ex+f} \qquad (a^2+d^2\neq 0)^{1}) \tag{26}$$

удаётся провести элементарными методами, хотя выполняемое при этом тождественное преобразование частично носит более или менее искусственный характер. Если $d\neq 0$, то следует сначала исключить целую часть из дроби, затем вместо остающейся дроби ввести величину, ей обратную:

$$\frac{ax^2+bx+c}{dx^2+ex+f}=\frac{a}{d}+\frac{1}{\dfrac{d(dx^2+ex+f)}{(bd-ae)x+(cd-af)}} \tag{27}$$

(оба выражения $bd-ae$ и $cd-af$ не обращаются одновременно в 0, иначе данная дробь не была бы несократимой). Дело сводится, таким образом, к исследованию дроби, у которой числитель — многочлен второй степени, а знаменатель — первой степени. То же самое мы имели бы, допустив сразу, что $d=0$.

Итак, рассмотрим дробь вида

$$y=\frac{mx^2+nx+p}{rx+s}, \tag{28}$$

причём допустим, что $m\neq 0$ и $r\neq 0$. Мы исключаем, таким образом, случаи, уже рассмотренные раньше (§§ 14 и 9). Вынося за скобки r и положив

$$\alpha=-\frac{s}{r},$$

мы можем расположить числитель по степеням $x-\alpha$ [2]):

$$y=\frac{m(x-\alpha)^2+(2m\alpha+n)(x-\alpha)+(m\alpha^2+n\alpha+p)}{r(x-\alpha)}.$$

Затем выделим целую часть, вынося $\frac{m}{r}$ за скобки:

$$y=\frac{m}{r}\left\{(x-\alpha)+\left(2\alpha+\frac{n}{m}\right)+\frac{\alpha^2+\frac{n}{m}\alpha+\frac{p}{m}}{x-\alpha}\right\}. \tag{29}$$

---

[1]) Запись $a^2+d^2\neq 0$ нужно понимать в её точном смысле: она означает, что хотя бы один из коэффициентов a и d отличен от нуля.

[2]) Для этого проще всего, введя новую переменную x_1, связанную с x равенствами $x_1=x-\alpha$, $x=x_1+\alpha$, расположить числитель по степеням x_1.

Величина $A=\alpha^2+\frac{n}{m}\alpha+\frac{p}{m}$ отлична от нуля[1]), но возможны два разных случая:

$$\text{I}\quad A<0$$

и

$$\text{II}\quad A>0.$$

В случае I мы придадим формуле (29) вид

$$y=\frac{2m\alpha+n}{r}+\frac{m\sqrt{|A|}}{r}\left\{\frac{x-\alpha}{\sqrt{|A|}}-\frac{\sqrt{|A|}}{x-\alpha}\right\},$$

откуда ясно, что график уравнения (28) в этом случае получается из графика уравнения

$$y=x-\frac{1}{x} \tag{I}$$

посредством последовательного выполнения следующих преобразований:

1) растяжения в $\sqrt{|A|}$ раз по направлению оси Ox,
2) растяжения в $\frac{m\sqrt{|A|}}{r}$ раз по направлению оси Oy,
3) перенесения параллельно оси Ox на отрезок $\alpha=-\frac{s}{r}$,
4) перенесения параллельно оси Oy на отрезок $\frac{2m\alpha+n}{r}$.

В случае же II, имея в виду, что $|A|=A$, мы напишем:

$$y=\frac{2m\alpha+n}{r}+\frac{m\sqrt{A}}{r}\left\{\frac{x-\alpha}{\sqrt{A}}+\frac{\sqrt{A}}{x-\alpha}\right\},$$

откуда ясно, что график (28) в этом случае посредством тех же преобразований, что и раньше, получается из графика

$$y=x+\frac{1}{x}. \tag{II}$$

Рассмотрим подробнее графики (I) и (II). Уравнению (I) можно придать вид

$$y=\frac{x^2-1}{x},$$

или

$$y=\frac{(x-1)(x+1)}{x}. \tag{30}$$

Функция (I) обращается в нуль в точках $x=1$ и $x=-1$ и теряет смысл (имеет «разрыв») в точке $x=0$ (рис. 19, a). Иссле-

---

[1]) Действительно, α есть корень знаменателя дроби (28); если бы было $m\alpha^2+n\alpha+p=0$, то α было бы также корнем числителя, и тогда дробь была бы сократима на $x-\alpha$.

дование её знака не представляет затруднений. Так как функция x — всюду возрастающая, а функция $\left(-\frac{1}{x}\right)$ — возрастающая всюду, кроме точки разрыва $x=0$, то из формулы (I) следует (§ 5, п. 3), что и о функции $x-\frac{1}{x}$ также можно утверждать, что она — возрастающая всюду, кроме точки разрыва $x=0$.

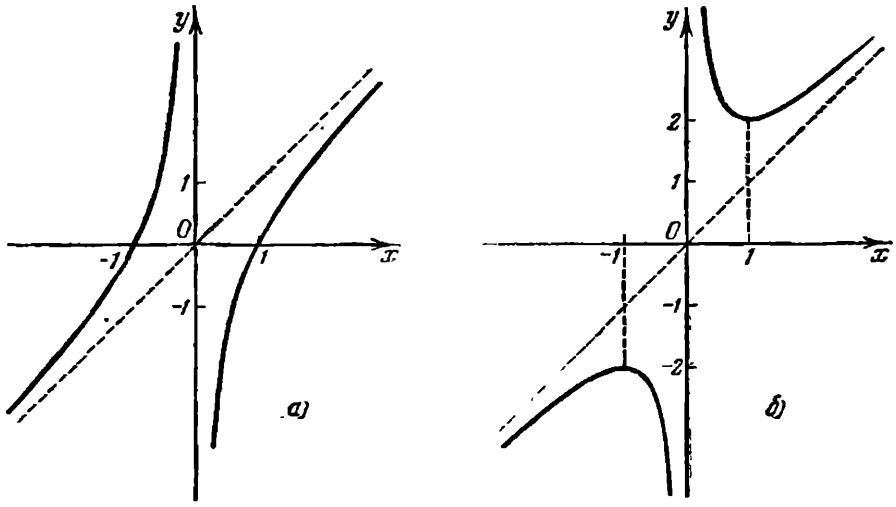


Рис. 19.

Что касается функции, заданной уравнением (II), то она также имеет разрыв в точке $x=0$, но в нуль нигде не обращается (рис. 19, *б*). Чтобы исследовать её изменяемость при $x>0$, придадим уравнению (II) вид

$$y=2+\left(\sqrt{x}-\frac{1}{\sqrt{x}}\right)^2. \qquad (31)$$

Эта запись показывает, что данная функция в точке $x=1$ принимает значение 2, при всех же прочих положительных значениях x принимает значения, большие чем 2; следовательно, она имеет минимум в точке $x=1$. Вместе с тем, так как в промежутке $1<x<\infty$ функции $\sqrt{x}$ и $\left(-\frac{1}{\sqrt{x}}\right)$ — возрастающие (§ 5, п.п. 8, 7, 6, 3), то и сумма их $\sqrt{x}-\frac{1}{\sqrt{x}}$ — также возрастающая; так как, кроме того, $\sqrt{x}-\frac{1}{\sqrt{x}}>0$, то возрастающей является (§ 5, п. 5) и функция $\left(\sqrt{x}-\frac{1}{\sqrt{x}}\right)^2$, следовательно, и функция, заданная уравнением (31), т. е. $x+\frac{1}{x}$. Аналогично устанавливается, что функ-

ция $x+\frac{1}{x}$ убывает в промежутке $0<x<1$: нужно только уравнению (II) придать вид

$$y=2+\left(\frac{1}{\sqrt{x}}-\sqrt{x}\right)^2. \tag{32}$$

Необходимость подробного исследования отрицательных значений x устраняется для функций (I) и (II) тем, что обе они — нечётные, откуда следует симметрия графика относительно центра O. Нужно отметить ещё присутствие в обоих случаях наклонной асимптоты $y=x$ (см. рис. 19, *а* и *б*); это значит, что при неограниченном возрастании x разность ординат $\left(x\pm\frac{1}{x}\right)-x$ неограниченно убывает.

Мы полностью выяснили поведение функций вида (28), т. е. функций вида (26) с ограничением $d=0$. Чтобы исследовать общий случай, достаточно констатировать, ссылаясь на формулу (27), что график функции общего вида (26) получается из графика функции вида (28) (в случае $bd-ac\neq 0$) или функции вида

$$y=ax^2+bx+c \qquad (a\neq 0)$$

(в случае $bd-ac=0$, $cd-af\neq 0$) посредством следующих преобразований:

1) нужно от графика данной функции перейти к графику функции, представляющей величину, ей обратную (см. § 2),

2) затем выполнить перенесение параллельно оси Oy на отрезок $\frac{a}{d}$.

Тем самым с помощью (§ 5, 1—9) определяется характер изменяемости рассматриваемой функции.

Исследование знака этой функции, а также определение тех точек, где она обращается в нуль или терпит разрыв, — всё это не представляет затруднений, если только степень числителя и знаменателя не выше чем 2.

Пример 1.

$$y=\frac{x^2-4}{x^2-2x-3}. \tag{33}$$

Запись

$$y=\frac{(x+2)(x-2)}{(x-3)(x+1)}$$

позволяет установить, что график пересекает ось Ox в точках $x=-2$ и $x=+2$ и имеет разрывы в точках $x=-1$ и $x=3$; она же даёт возможность исследовать знак y. Кроме того, из записи

$$y=\frac{1-\frac{4}{x^2}}{1-\frac{2}{x}-\frac{3}{x^2}}$$

видно, что при неограниченном увеличении $|x|$ величина y приближается к 1: прямая $y=1$ является асимптотой. Сделав дополнительно к этому 2—3 точечные подстановки, мы намечаем график, изображённый на рис. 20.

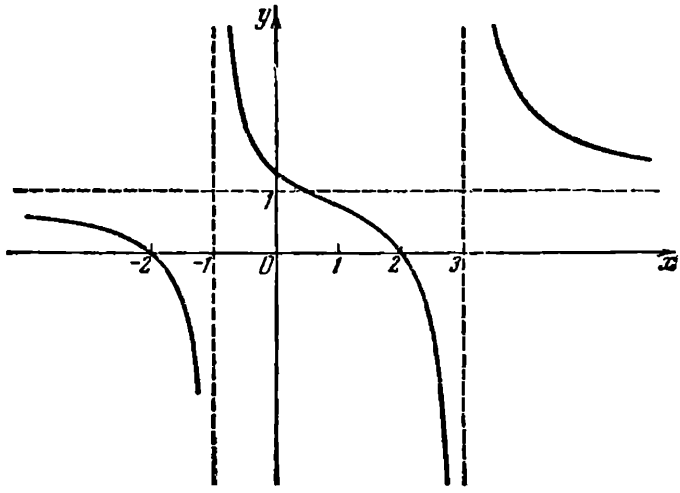


Рис. 20.

Однако для того, чтобы убедиться в том, что функция (33) всюду, кроме двух точек разрыва, убывает, нужно прибегнуть к изложенной выше теории. Следуя ей, мы приходим к записи

$$y=1+\cfrac{2}{-1+\left(x-\frac{1}{2}\right)+\cfrac{\left(-\frac{15}{4}\right)}{x-\frac{1}{2}}},$$

и тогда для получения нужного заключения остаётся непосредственно сослаться на теоремы § 5.

Пример 2.

$$y=\frac{x^2+x+1}{x^2-x+1}.$$

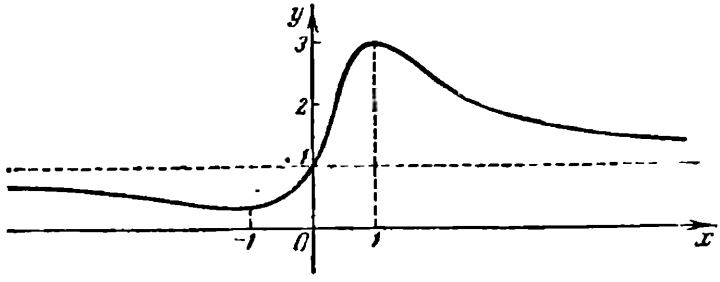


Рис. 21.

Числитель и знаменатель всегда положительны; следовательно, всегда $y>0$. Прямая $y=1$ — асимптота, как и в примере 1. Изложенная теория

приводит к записи

$$y=1+\frac{2}{-1+\left(x+\frac{1}{x}\right)}.$$

Функция $x+\frac{1}{x}$ убывает при $|x|<1$ (кроме точки разрыва) и возрастает при $|x|>1$; отсюда следует, что функция y возрастает при $|x|<1$ и убывает при $|x|>1$, но никаких разрывов у неё уже нет (рис. 21).

Пример 3.

$$y=\frac{1}{x^2+1}.$$

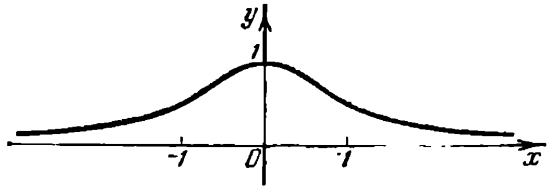


Рис. 22.

Знак функции всегда положителен; характер изменяемости следует непосредственно из теорем § 5 (рис. 22).

Пример 4.

$$y=\frac{x^2-1}{x^2+1}.$$

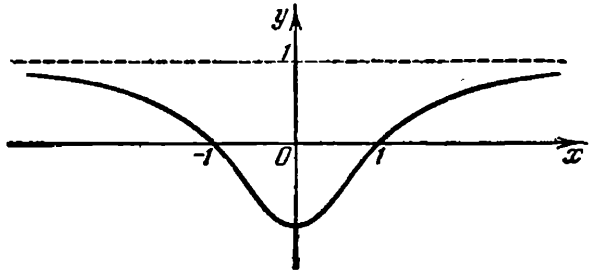


Рис. 23.

Знак функции меняется при $x=\pm 1$ (рис. 23); для исследования изменяемости достаточно согласно теории выделить целую часть:

$$y=1-\frac{2}{x^2+1}.$$

§ 16. Дробные рациональные функции (общий случай)

Всякая дробная рациональная функция $y=R(x)$ может быть представлена в виде отношения двух многочленов, не имеющих общих корней,

$$R(x)=\frac{P(x)}{Q(x)},$$

где

$$P(x)=ax^n+bx^{n-1}+\ldots+kx+l=$$
$$=a(x-\alpha)^{\lambda}(x-\beta)^{\mu}\ldots(x^2+px+q)^{\rho}\ldots,$$
$$Q(x)=a'x^{n'}+b'x^{n'-1}+\ldots+k'x+l'=$$
$$=a'(x-\alpha')^{\lambda'}(x-\beta')^{\mu'}\ldots(x^2+p'x+q')^{\rho'}\ldots$$

Возникающие при разложении на множители трёхчлены второй степени с мнимыми корнями x^2+px+q, $x^2+p'x+q'$, ... всегда сохраняют положительный знак. Все корни числителя и знаменателя предполагаются различными. Старшие коэффициенты a и a' (несущественно ограничивая общность) можно для простоты принять равными единице. Итак,

$$y=R(x)=\frac{(x-\alpha)^{\lambda}(x-\beta)^{\mu}\ldots(x^2+px+q)^{\rho}\ldots}{(x-\alpha')^{\lambda'}(x-\beta')^{\mu'}\ldots(x^2+p'x+q')^{\rho'}\ldots} \qquad (34)$$

Представим себе, что числа α, β, ..., α', β', ..., являющиеся действительными корнями числителя и знаменателя, будучи расположены в порядке возрастания, разбивают ось Ox на некоторое число промежутков. Взяв какое-нибудь значение x в одном из промежутков, легко сосчитаем, с учётом кратности, число множителей в числителе и знаменателе дроби, имеющих отрицательный знак при этом значении x: знак функции будет тот или иной в зависимости от чётности этого числа. Названное число множителей равно числу действительных корней числителя и знаменателя нечётной кратности, превышающих рассматриваемое значение x: оно, очевидно, не зависит от выбора значения x в данном промежутке и зависит, таким образом, только от промежутка. Исходя из этого соображения, можно сразу разметить знаки, соответствующие промежуткам, представляя себе, что x убывает от $+\infty$ до $-\infty$, и учитывая, что знак функции меняется при прохождении x через всякий действительный корень нечётной кратности в числителе и знаменателе.

Если a и a' не равны между собой, то перед дробью в формуле (34) появляется числовой коэффициент: будучи положительным, он не влияет на знак функции (вызывая растяжение графика функции в направлении оси Oy); будучи отрицательным, даёт противоположное распределение знаков (к растяжению присоединяется отражение относительно оси Ox).

Выясним подробнее расположение графика в окрестности такой точки, в которой функция меняет знак. Объединяя вместе случаи, когда такая точка оказывается корнем числителя или корнем знаменателя дроби в формуле (34), придадим этой формуле вид

$$y = R(x) = (x - \xi)^{\sigma} R_1(x),$$

где ξ — рассматриваемый корень одного из многочленов, $|\sigma|$ — его кратность, $R_1(x)$ — произведение всех прочих множителей (с положительными и отрицательными степенями).

Если ξ — корень числителя, т. е. $\sigma > 0$, то функция $R(x)$ при $x = \xi$ принимает значение нуль, и получается точка графика на оси Ox; график в этой точке касается оси или пересекает [1]) ось

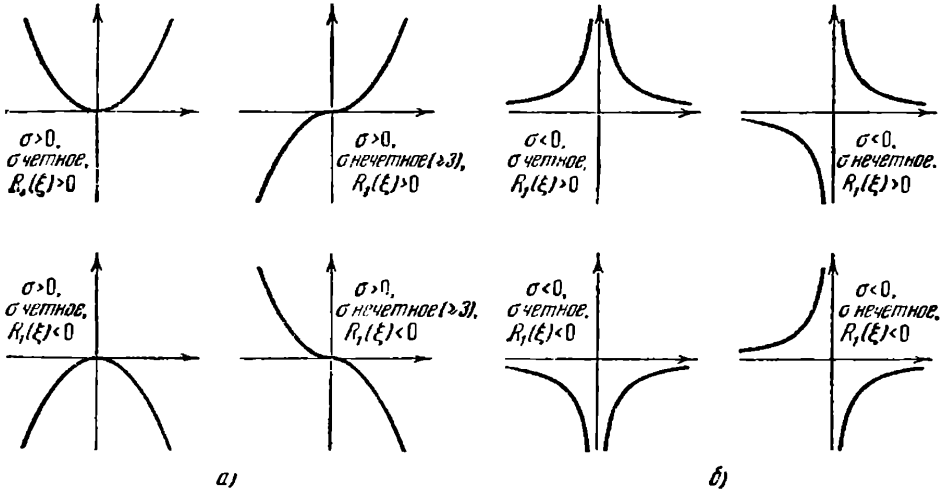


Рис. 24.

в зависимости от того, будет ли σ чётным или нечётным. Если ещё обратить внимание на то, что возможен тот или иной знак $R_1(\xi)$, то становится ясным, что имеет место одно из четырёх расположений, схематически указанных на рис. 24, *а*.

В случае, если $\sigma < 0$, т. е. ξ есть корень знаменателя (и, следовательно, «точка разрыва» функции $R(x)$), соответствующие четыре расположения показаны на рис. 24, *б*.

Чтобы выяснить вопрос о «поведении кривой на бесконечности», достаточно разделить числитель и знаменатель данной дроби на x^N, где N — наибольшее из чисел n и n' (степеней числителя и знаменателя). Оказывается, что при неограниченном возрастании $|x|$:

1) если $n < n'$, кривая асимптотически приближается к оси Ox,

[1]) Пересекает без касания, если $\sigma = 1$; пересекает с касанием, если $\sigma = 3, 5, \ldots$

2) если $n=n'$, то она имеет горизонтальную асимптоту

$$y=\frac{a}{a'},$$

3) если $n>n'$, то $|y|$ неограниченно возрастает при увеличении $|x|$; при этом в случае $n\geqslant n'+2$ прямолинейных асимптот нет, а в случае $n=n'+1$ имеется наклонная асимптота, уравнение которой легко составляется посредством исключения целой части из дроби. Например, кривая

$$y=\frac{2x^3+3x^2}{x^2+x+1},$$

как видно из записи

$$y=2x+1-\frac{3x+1}{x^2+x+1},$$

имеет асимптоту $y=2x+1$.

§ 17. Алгебраические иррациональные функции

Целой рациональной функцией (многочленом) от двух переменных x и y называется такая функция, значения которой могут быть получены из значений этих переменных и из постоянных чисел посредством не более чем трёх операций — сложения, вычитания и умножения.

Если $P(x, y)$ — такая функция, то уравнение

$$P(x, y)=0 \tag{35}$$

(называемое *алгебраическим*) определяет алгебраическую функциональную зависимость между x и y.

Предположим, что некоторая функция, однозначная или многозначная,

$$y=f(x) \tag{36}$$

в некотором промежутке удовлетворяет уравнению вида (35) тождественно относительно x:

$$P(x, f(x))\equiv 0.$$

Тогда функция (36) называется *алгебраической*. Например, функция $y=\sqrt{1-x^2}$ — алгебраическая, так как в промежутке $-1\leqslant y\leqslant 1$ удовлетворяет уравнению $x^2+y^2=1$ [1]).

Всякая рациональная функция (в том числе и *целая*) *является алгебраической.*

Действительно, из соотношения

$$y=\frac{P(x)}{Q(x)},$$

[1]) См., впрочем, стр. 214.

где $P(x)$ и $Q(x)$ — многочлены (целые рациональные функции) относительно x, следует соотношение

$$Q(x)\,y - P(x) = 0,$$

причём его левая часть есть целая рациональная функция относительно переменных x и y.

Многочлены от двух переменных x и y, как и многочлены от одной переменной, классифицируются по степеням. Необходимо только указывать, имеются ли в виду степени многочлена относительно переменной x или относительно y, или относительно совокупности переменных x и y; в последнем случае

1) степенью отдельного члена называется сумма степеней x и y, в него входящих,

2) степенью многочлена называется наибольшая из степеней отдельных членов.

Например, многочлен

$$P(x,\ y) = x^2 + 3x^5y^3 - y^4$$

— степени 5 относительно x, степени 4 относительно y и степени 8 относительно совокупности x и y.

Если левая часть уравнения (35) — степени n относительно совокупности переменных x и y, то говорят, что само уравнение, а также его график — порядка n.

Из предыдущего ясно, что всякая рациональная функция $y = \frac{P(x)}{Q(x)}$ удовлетворяет алгебраическому уравнению первой степени относительно y. Но это не означает, что всякое уравнение, которому удовлетворяет рациональная функция, непременно должно быть первой степени относительно y. Например, рациональная функция $y = x$ удовлетворяет уравнению второй степени относительно y

$$y^2 = x^2, \tag{37}$$

а также уравнению третьей степени относительно y

$$y^3 = x^3; \tag{38}$$

при этом уравнению (37) удовлетворяет также другая рациональная функция $y = -x$; но уравнению (38) не удовлетворяет никакая функция, кроме $y = x$.

Те алгебраические функции, которые не являются рациональными, называются *иррациональными*.

В качестве простейших примеров иррациональных алгебраических функций можно указать

$$y = \sqrt{x}$$

или

$$y = \sqrt{x^2 + 1}\ ^{1}).$$

---

[1]) В этом параграфе и дальше, если не сделано оговорки, радикалы понимаются в алгебраическом смысле: это значит, что в случае чётного показателя при радикале подразумевается двойной знак.

Если мы хотим доказать иррациональность данной функции, то нам нужно установить, что её **нельзя** представить в виде отношения двух многочленов. Как всякое доказательство невозможности, оно связано с известными трудностями. Читателю можно порекомендовать попытаться найти доказательство иррациональности двух приведённых выше функций, строя его по образцу известного доказательства иррациональности числа $\sqrt{2}$ и опираясь при этом на лемму: *если квадрат многочлена делится на* x (или на x^2+1), *то и сам многочлен делится на* x (или на x^2+1).

§ 18. Примеры исследования алгебраических функций

Рассмотрим несколько характерных примеров.

Пример 1. Функция

$$y=\sqrt{x} \qquad (x\geqslant 0) \tag{39}$$

является обратной по отношению к функции $y=x^2$. В самом деле, она удовлетворяет уравнению

$$y^2-x=0. \tag{40}$$

Поменяв местами x и y (см. § 4), получаем уравнение

$$y-x^2=0, \tag{41}$$

которое нами уже было исследовано (§ 7). График уравнения (41) — парабола с осью Oy и вершиной O, расположенная в верхней полу-

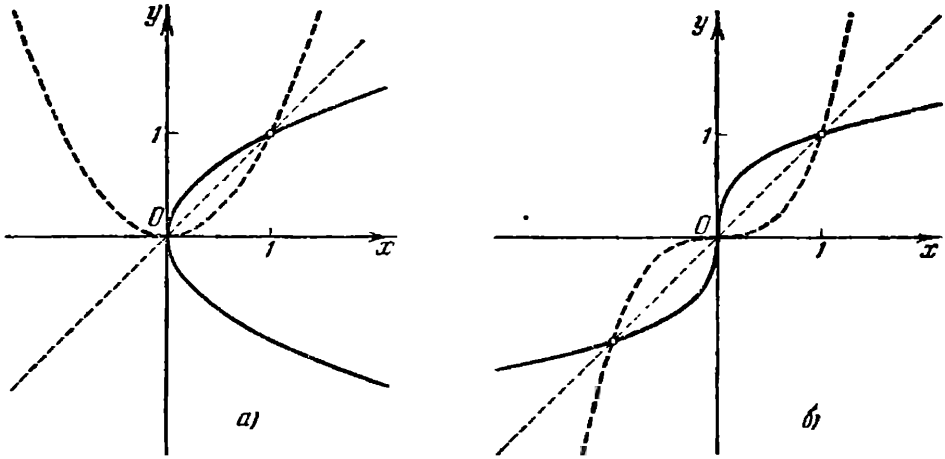


Рис. 25.

плоскости (см. рис. 13, *б*). График уравнения (40) получается из графика (41) посредством преобразования симметрии относительно биссектрисы $y=x$: это — парабола с осью Ox, расположенная в правой полуплоскости (рис. 25, *а*).

График (41) в вершине O имеет горизонтальную касательную (ось Ox); график (40) в той же вершине имеет вертикальную касательную (ось Oy).

Уравнение (40) определяет двузначную функцию y, поэтому в формуле (39) естественно понимать радикал в алгебраическом смысле — с двойным знаком. Если взять только знак «$+$» (арифметическое значение радикала), то в качестве графика получим «половину» параболы, лежащую в верхней полуплоскости; аналогично для знака «$-$».

Пример 2. Подобным же образом функция

$$y=\sqrt[3]{x} \tag{42}$$

— обратная по отношению к функции

$$y=x^3. \tag{43}$$

График (43) — кубическая парабола с вершиной O, проходящая в первом и третьем квадрантах и в точке O касающаяся оси Ox (см. § 7; рис. 13, *в*). График (42) — симметричная ей (относительно биссектрисы $y=x$) парабола; у неё та же вершина, и она проходит через те же квадранты, но в вершине O имеет вертикальную касательную Oy (рис. 25, *б*).

Обе функции (43) и (42) однозначны и определены для всех значений x.

Пример 3. Вообще при любом целом положительном n функция

$$y=\sqrt[n]{x} \tag{44}$$

является обратной по отношению к функции

$$y=x^n. \tag{45}$$

Графики функций (44), получающиеся по симметрии из графиков функций (45), имеют в точке O вертикальную касательную и примыкают к ней тем теснее, чем больше n.

Расположение кривой при любом чётном n — такое же, как при $n=2$; при любом нечётном n — как при $n=3$.

Пример 4. Рассмотрим произвольную рациональную степень

$$y=x^{\frac{p}{q}} \quad (p\neq q), \tag{46}$$

или

$$y=\sqrt[q]{x^p}, \tag{47}$$

где p и q — целые положительные числа, не имеющие общих множителей. Эта функция удовлетворяет уравнению

$$x^p-y^q=0. \tag{48}$$

Ограничиваясь рассмотрением положительных значений x и y, заметим, что отношение

$$\frac{y}{x} = x^{\frac{p}{q}-1}$$

при неограниченном приближении x к нулю само либо неограниченно убывает, либо, напротив, неограниченно возрастает, — смотря по тому, будет ли $\frac{p}{q} > 1$ или $\frac{p}{q} < 1$; геометрически это означает, что, приближаясь к началу координат в пределах первого квадранта, график либо примыкает к оси Ox (рис. 26, *а*), либо — к оси

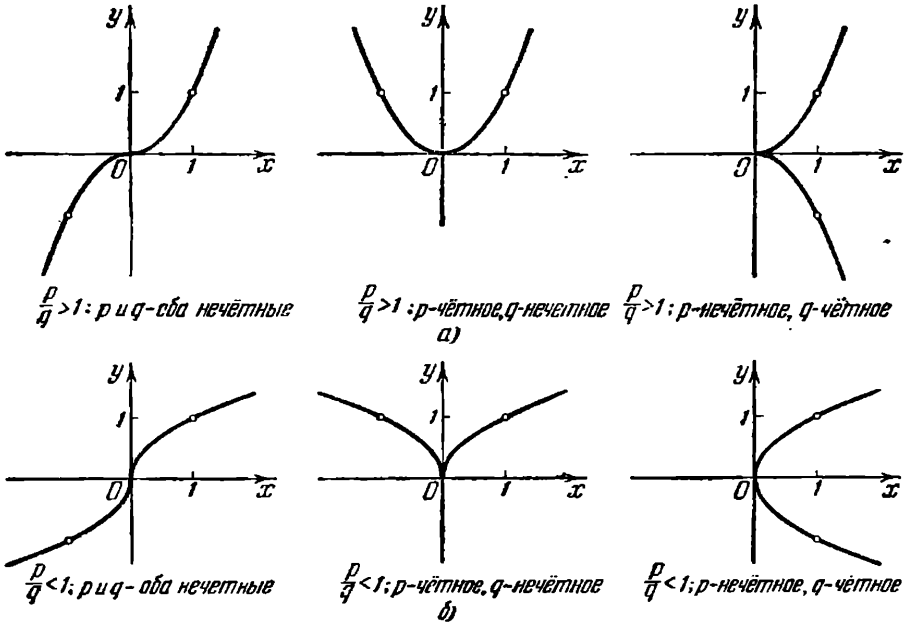


Рис. 26.

Oy (рис. 26, *б*) («касается» той или другой), смотря по тому, будет ли показатель $\frac{p}{q}$ больше или меньше единицы. При $x > 0$ функция $x^{\frac{p}{q}}$ возрастает и тем быстрее, чем больше степень $\frac{p}{q}$.

Что касается поведения графика в остальных квадрантах, то, глядя на уравнение (48), легко понять, что

1) если p и q — оба нечётные, то имеется симметрия относительно центра O, так что график лежит в 1-м и 3-м квадрантах,

2) если p чётное, а q нечётное, то имеется симметрия относительно оси Oy, так что график лежит в 1-м и 2-м квадрантах,

3) если p — нечётное, а q — чётное, то имеется симметрия относительно оси Ox, так что график лежит в 1-м и 4-м квадрантах.

Возможные случаи схематически изображены на рис. 26.

Следует обратить внимание на наличие так называемой «точки возврата» в начале координат O в том случае, когда числа p и q — различной чётности, причём нечётное больше чётного.

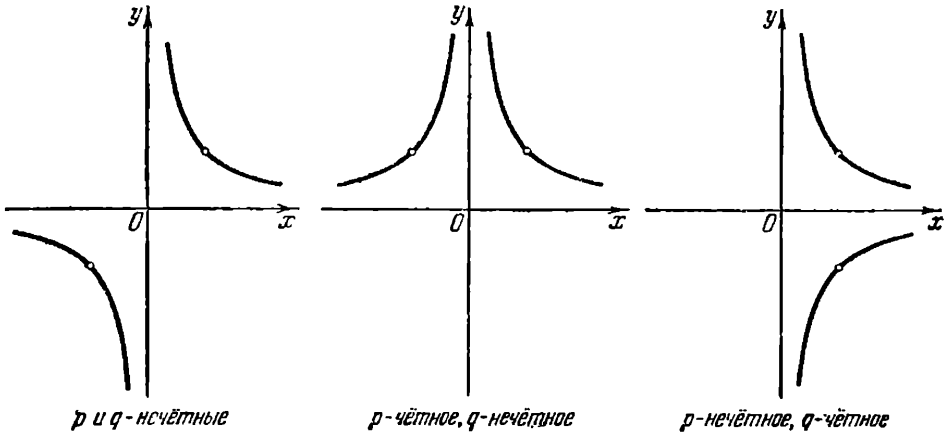
Пример 5. Функции

$$y=x^{-\frac{p}{q}},$$

или

$$y=\frac{1}{\sqrt[q]{x^p}}$$

(при целых положительных p и q) представляют собой величины, обратные по отношению к функциям, рассмотренным в предыду-



p и q - нечётные *p - чётное, q - нечётное* *p - нечётное, q - чётное*

Рис. 27.

щем примере. Ограничимся (ссылаясь на § 2, п. 6 и § 5, п. 7) приведением рис. 27, показывающего схематически возможные расположения графиков в окрестности точки $x=0$.

Пример 6. Функция

$$y=\sqrt{-x} \qquad (x\leqslant 0) \tag{49}$$

имеет график, симметричный графику функции (40) (см. рис. 25, *а*) относительно оси Oy.

Пример 7. Функция

$$y=\sqrt{1-x^2} \qquad (|x|\leqslant 1) \tag{50}$$

удовлетворяет алгебраическому уравнению второго порядка

$$x^2+y^2=1. \tag{51}$$

График этого уравнения — окружность радиуса 1 с центром в начале координат O (рис. 28); действительно, из уравнения видно, что расстояние точки $M(x, y)$ от начала O равно 1.

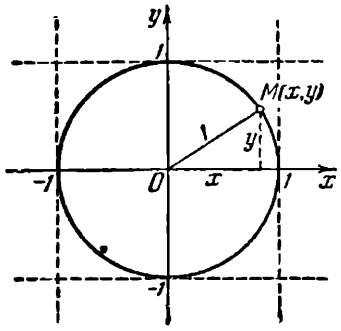


Рис. 28.

При растяжении в r раз по направлениям осей Ox и Oy мы получим уравнение

$$x^2+y^2=r^2, \qquad (52)$$

график которого — окружность радиуса r с центром O. Производя теперь перенесения параллельно оси Ox на отрезок a и параллельно оси Oy на отрезок b, приходим к уравнению

$$(x-a)^2+(y-b)^2=r^2, \qquad (53)$$

представляющему геометрически окружность радиуса r с центром (a, b).

Пример 8. При растяжении в p раз по направлению оси Ox и в q раз по направлению оси Oy уравнение (51) примет вид

$$\frac{x^2}{p^2}+\frac{y^2}{q^2}=1, \qquad (54)$$

или

$$y=\frac{q}{p}\sqrt{p^2-x^2} \qquad (|x|\leqslant p);$$

графиком этого уравнения является эллипс с полуосями p и q (рис. 29). Уравнение (54) называется *каноническим уравнением эллипса*. При $p=q=r$ получается окружность (52) радиуса r с центром O.

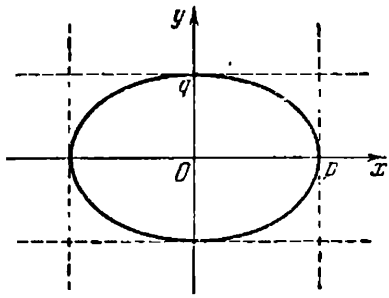


Рис. 29.

Пример 9. Преобразования IV а и IV б, рассмотренные в § 3, будучи применены соответственно к уравнениям

$$\text{а)}\quad x^2-y^2=1, \qquad y=\sqrt{x^2-1}\ (|x|\geqslant 1), \qquad (55)$$

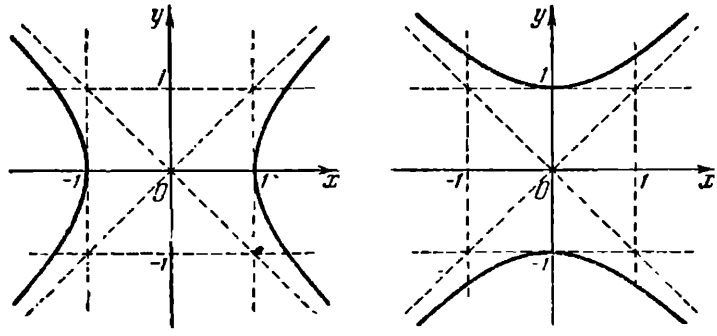
$$\text{б)}\quad x^2-y^2=-1, \qquad y=\sqrt{x^2+1}, \qquad (56)$$

в обоих случаях дают уравнение

$$2xy=1,$$

и остаётся ещё сделать растяжение в $\sqrt{2}$ раз по направлениям обеих осей, чтобы убедиться в том, что графики данных алгебраи-

ческих иррациональных функций — гиперболы (см. § 13, рис. 17). Расположение их указано на рис. 30, *а* и *б*.



а) *б*)

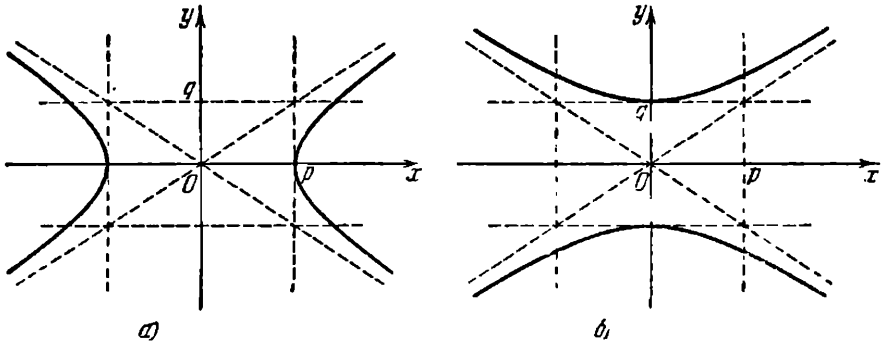
Рис. 30.

Пример 10. Растяжение в p раз по направлению оси Ox и в q раз по направлению оси Oy от уравнений а) и б) примера 9 приводит к уравнениям

$$\text{а')}\ \frac{x^2}{p^2}-\frac{y^2}{q^2}=1, \qquad y=\frac{q}{p}\sqrt{x^2-p^2}\quad (|x|\geqslant p), \tag{57}$$

$$\text{б')}\ \frac{x^2}{p^2}-\frac{y^2}{q^2}=-1, \quad y=\frac{q}{p}\sqrt{x^2+p^2} \tag{58}$$

(графики — сопряжённые между собой гиперболы с полуосями p и q (рис. 31, *а* и *б*)).



а) *б*)

Рис. 31.

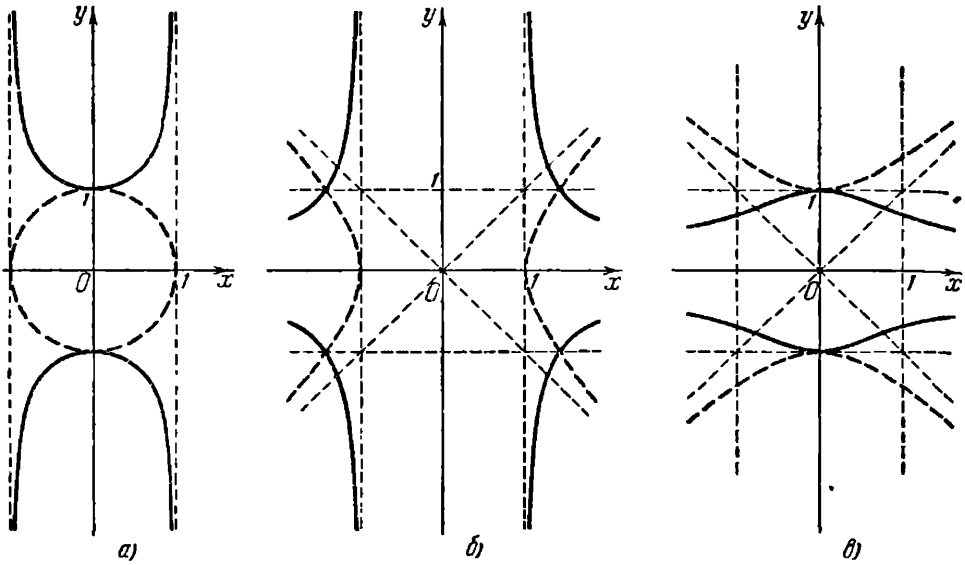
Пример 11. Функции

$$\left.\begin{aligned} &\text{а)}\ y=\frac{1}{\sqrt{1-x^2}} \quad (|x|<1),\\ &\text{б)}\ y=\frac{1}{\sqrt{x^2-1}} \quad (|x|>1),\\ &\text{в)}\ y=\frac{1}{\sqrt{x^2+1}}, \end{aligned}\right\} \tag{59}$$

удовлетворяющие соответственно уравнениям четвёртого порядка

$$x^2y^2+1=y^2,\quad x^2y^2=y^2+1,\quad x^2y^2+y^2=1,$$

представляют собой величины, обратные функциям (50), (55) и (56). Этим определяется характер их изменения. На рис. 32, *а*, *б*, *в* изо-



а) *б)* *в)*

Рис. 32.

бражены соответственно графики (59) а), б), в) (сплошные линии) и для сравнения графики (50), (55) и (56) (пунктирные линии).

Пример 12.

$$\sqrt{x}+\sqrt{y}=1. \tag{60}$$

Освободившись от радикалов, приходим к уравнению второго порядка

$$(x-y)^2+1=2(x+y). \tag{60'}$$

Оно определяет двузначную алгебраическую функцию

$$y=(1-\sqrt{x})^2 \qquad (x\geqslant 0). \tag{61}$$

График имеет осью симметрии биссектрису $y=x$, так как уравнение (60) или (60') не изменяется при взаимной перестановке x и y. Он целиком лежит в первом квадранте и на самих осях имеет только точки (0, 1) и (1, 0). Поворот осей на -45° (см. § 3) приводит уравнение (60') к виду

$$y'-\frac{1}{2\sqrt{2}}=\frac{x'^2}{\sqrt{2}}; \tag{62}$$

отсюда ясно, что график данного уравнения — парабола, в точках (0, 1) и (1, 0) касающаяся координатных осей [1]) (рис. 33).

Пример 13. Четырёхзначная функция

$$y=\sqrt{x+1}+\sqrt{x} \quad (x\geqslant 0) \quad (63)$$

удовлетворяет уравнению четвёртого порядка

$$(y^2-1)^2=4xy^2, \quad (64)$$

которое легко получить, положив

$$u=\sqrt{x+1},\ v=\sqrt{x},\ y=u+v,$$

и затем исключая u и v. В расположении графика легко отдать себе отчёт, или принимая y за независимое переменное

$$x=\frac{(y^2-1)^2}{4y^2},$$

или путём «сложения» (см. § 2) параболических графиков $y=\sqrt{x}$ и $y=\sqrt{x+1}$.

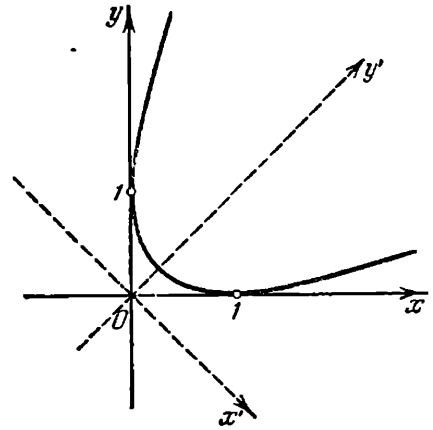


Рис. 33.

«Расщепим» одну четырёхзначную функцию (63) на четыре однозначные, причём введём радикалы в арифметическом смысле:

$$y_1=\sqrt{x+1}+\sqrt{x},$$

$$y_2=\sqrt{x+1}-\sqrt{x}=$$

$$=\frac{1}{\sqrt{x+1}+\sqrt{x}}=\frac{1}{y_1},$$

$$y_3=-y_2,\quad y_4=-y_1.$$

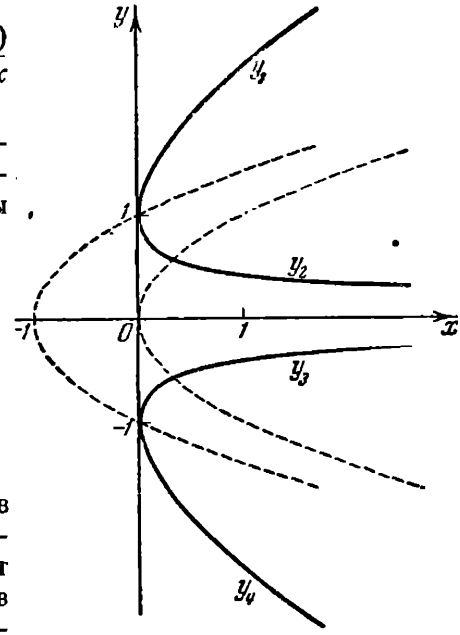


Рис. 34.

Графики y_1 и y_2 оба лежат в первом квадранте, причём y_1 возрастает (§ 5, п. 3), а y_2 убывает (§ 5, п. 7); они смыкаются в точке (0, 1), где имеется вертикальная касательная (рис. 34). При

[1]) В самом деле, уравнение (62) имеет графиком параболу, получающуюся из параболы $y=x^2$ посредством растяжения в $\sqrt[4]{2}$ раз по направлению оси Ox и перемещения параллельно оси Oy на отрезок $\frac{1}{2\sqrt{2}}$.

неограниченном возрастании x функция y_1 неограниченно возрастает, хотя и довольно медленно [1]); напротив, функция y_2 неограниченно стремится к нулю [2]).

Графики y_3 и y_4 симметричны графикам y_2 и y_1 относительно оси Ox.

Пример 14. По уравнению четвёртого порядка

$$x^4+y^4=x^2+y^2 \tag{65}$$

(см. § 1, стр. 16) можно судить о наличии нескольких симметрий (§ 3): относительно оси Ox, относительно оси Oy (значит, и относительно центра O), а также относительно биссектрисы $y=x$. Поэтому при исследовании графика уравнения (65) достаточно ограничиться рассмотрением «восьмушки» плоскости, определяемой неравенствами

$$0\leqslant x\leqslant y.$$

При таких условиях в решении уравнения относительно y

$$y=\sqrt{\frac{1}{2}+\sqrt{\frac{1}{4}+x^2-x^4}} \tag{66}$$

(предполагая, что оба радикала взяты в арифметическом смысле) достаточно ограничиться рассмотрением значений x, не превосходящих 1 (при $x>1$ мы получили бы $\sqrt{\frac{1}{4}+x^2-x^4}<\frac{1}{2}$, $y<1$, что противоречит условию $x\leqslant y$); знак минус перед внутренним радикалом при $x>0$ брать и подавно излишне. При возрастании x от 0 до 1 изменение функции y зависит, очевидно, от изменения выражения

$$z=\frac{1}{4}+x^2-x^4=\frac{1}{2}-\left(x^2-\frac{1}{2}\right)^2$$

(см. § 9). Величина z не превышает $\frac{1}{2}$ и равна $\frac{1}{2}$ только при $x=\frac{1}{\sqrt{2}}$; значит, y не превышает $\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}$ и принимает это наибольшее возможное значение только при $x=\frac{1}{\sqrt{2}}$. При возра-

---

[1]) Точнее говорит об этом соотношение

$$y_1-2\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}<\frac{1}{2\sqrt{x}}.$$

[2]) Именно,

$$y_2<\frac{1}{2\sqrt{x}}.$$

стании x от 0 до $\frac{1}{\sqrt{2}}$ величина $\left(x^2-\frac{1}{2}\right)^2$ убывает от $\frac{1}{4}$ до 0; значит, z возрастает от $\frac{1}{4}$ до $\frac{1}{2}$, а y возрастает от 1 до $\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}$; при возрастании, далее, x от $\frac{1}{\sqrt{2}}$ до 1 величина $\left(x^2-\frac{1}{2}\right)^2$ возрастает от 0 до $\frac{1}{4}$; значит, z убывает от $\frac{1}{2}$ до $\frac{1}{4}$, а y убывает от $\sqrt{\frac{1}{2}+\sqrt{\frac{1}{2}}}$ до 1.

Следует обратить внимание на то, что при $x=0$ (но только при этом значении) в формуле (66) можно взять знак минус перед внутренним радикалом. Это даст «изолированную точку» в начале O. Весь график изображён на рис. 35.

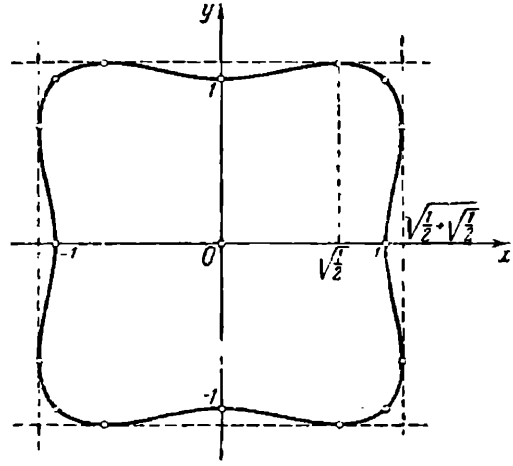


Рис. 35.

Примечание. Если уравнение $y=f(x)$ таково, что правая часть его «составлена» из x и постоянных чисел с помощью лишь рациональных операций и извлечения радикалов целых степеней (в конечном числе), то y есть элементарная, и притом алгебраическая, функция переменной x. Действительно, «избавляясь» от радикалов, уравнению указанного вида всегда можно придать вид (35)[1]).

Напротив, если мы исходим из уравнения вида (35), то, допуская, что существует функция $y=f(x)$, ему удовлетворяющая[2]), мы можем утверждать, что эта функция — алгебраическая, но не всегда верно, что она — элементарная. В самом деле, например, уравнение

$$y^5+y=x$$

невозможно решить в радикалах относительно y (при произвольном буквенном параметре x); тем не менее легко доказывается, что это уравнение определяет y однозначно как функцию x: при возрастании x от $-\infty$ до $+\infty$ левая часть уравнения также возрастает от $-\infty$ до $+\infty$ (рис. 36) и, следовательно, при одном и только одном значении y примет любое наперёд заданное значение x (см. ниже § 52, теорема Больцано). Вполне понятно, что здесь идёт речь о функции как о «соответствии».

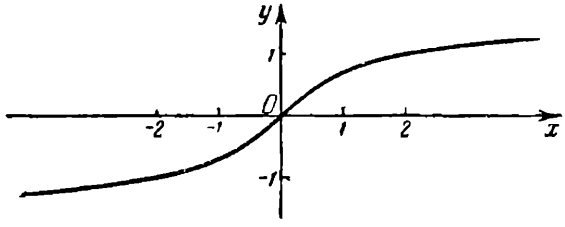


Рис. 36.

[1]) Чтобы «избавиться» от радикалов, нужно принять их за новые переменные и затем исключить (см. Э. э. м., кн. 1, А. И. Узков, Векторные пространства и линейные преобразования, § 17).

[2]) Заметим, что в примере $x^2+y^2+1=0$ не существует никакой функции действительной переменной, которая удовлетворяла бы уравнению.

§ 19. Элементарные трансцендентные функции

Термин «трансцендентная» функция имеет вполне точный смысл: функция $y=f(x)$ называется *трансцендентной*, если она не удовлетворяет никакому алгебраическому уравнению

$$P(x, y)=0,$$

где $P(x, y)$ обозначает алгебраический многочлен (целую рациональную функцию) относительно переменных x и y. В курсе математики средней школы из числа трансцендентных функций систематически изучаются: показательные, тригонометрические (круговые), а также им обратные; но учащемуся приходится встречаться и с различными их комбинациями.

В высшей математике рассматриваются и изучаются трансцендентные функции самых разнообразных типов.

«Трансцендентный» обозначает буквально — «превосходящий» (подразумевается, по Эйлеру, превосходящий силу алгебраических методов, что соответствует в точности приведённому выше определению). Доказательства трансцендентности функций относятся к числу «доказательств невозможности», они строятся по схеме «от противного» и требуют привлечения разного рода искусственных приёмов.

§ 20. Показательная функция

Показательной, или *экспоненциальной* [1]), функцией называют всякую функцию вида

$$y=a^x \qquad (a>0). \tag{67}$$

Символу a^x даётся «расчленённое» определение.

1°. Если x равно целому положительному числу n, то a^x следует понимать как результат повторного умножения:

$$a^n=\underbrace{aa\ldots a}_{n \text{ раз}}. \tag{68}$$

2°. Если x есть положительное дробное рациональное число $\frac{p}{q}$, то, возводя основание a в степень x согласно формуле

$$a^{\frac{p}{q}}=\sqrt[q]{a^p}\ , \tag{69}$$

берут арифметическое значение радикала.

---

[1]) «Экспонент» (по-латыни) означает показатель.

3°. Если x — отрицательное число $\left(-\frac{p}{q}\right)$, то полагают

$$a^{-\frac{p}{q}}=\frac{1}{a^{\frac{p}{q}}}; \tag{70}$$

наконец,

$$a^0=1. \tag{71}$$

4°. Для иррациональных значений x показательная функция a^x определяется «по непрерывности» (см. ниже § 51).

Примечание. В этом параграфе сформулированы важнейшие свойства показательной функции, но доказательства приведены ниже лишь для случая рациональных показателей. По поводу иррациональных показателей см. § 51.

Из определения следует, что при всех значениях независимой переменной x показательная функция принимает положительные значения. В частности, она ни при каких значениях x не обращается в нуль. Её график весь располагается выше оси Ox.

Рассматривать показательную функцию при отрицательном основании a не приходится по той причине, что раз a отрицательно, переменной x нельзя давать дробных значений вида $\frac{2m+1}{2n}$ (где m и n целые), не говоря уже о значениях иррациональных (см. стр. 510).

Случаи, когда $a=0$ или $a=1$, не представляют никакого интереса.

Показательная функция при основании a, большем единицы, — возрастающая; при основании a, меньшем единицы, — убывающая.

Пусть даны два рациональных числа x' и x'', причём $x'<x''$. Допустим, что они приведены к общему знаменателю:

$$x'=\frac{p'}{q},\quad x''=\frac{p''}{q},\quad p'<p''.$$

Тогда из неравенства $a>1$ следует неравенство $a^{p'}<a^{p''}$ и, далее, после извлечения арифметических корней степени q [1]),

$$a^{\frac{p'}{q}}<a^{\frac{p''}{q}},\ \text{т. е.}\quad a^{x'}<a^{x''}.$$

Таким же образом из неравенства $0<a<1$ мы получили бы $a^{p'}>a^{p''}$, $a^{\frac{p'}{q}}>a^{\frac{p''}{q}}$, т. е. $a^{x'}>a^{x''}$.

Чтобы выяснить взаимное расположение графиков двух показательных функций a^x и b^x $(0<a<b)$, обратим прежде всего вни-

[1]) Функция $x^{\frac{1}{q}}$ — возрастающая (при $x>0$), см. § 18, пример 3.

мание на то, что они пересекаются в точке (0, 1), (так как $a^0 = 1$ и $b^0 = 1$), и затем установим следующее: если $0 < a < b$, то при $x > 0$ имеет место неравенство $a^x < b^x$, а при $x < 0$ — неравенство $a^x > b^x$ [1]).

Заметим, что график функции $\left(\frac{1}{a}\right)^x$ симметричен графику a^x относительно оси Oy.

Действительно, мы имеем

$$\left(\frac{1}{a}\right)^x = a^{-x}$$

и остаётся сослаться на I б, § 3.

На рис. 37 даны графики нескольких показательных функций при различных основаниях.

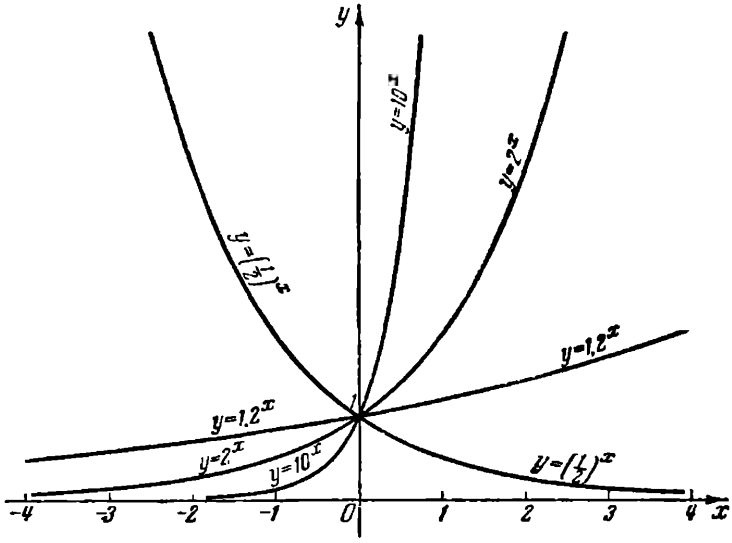


Рис. 37.

При неограниченном возрастании x показательная функция $y = a^x$ (если $a > 1$) *также неограниченно возрастает и притом очень быстро — быстрее чем любая степень x.*

Принимая во внимание, что функция a^x — возрастающая, достаточно ограничиться рассмотрением целых значений переменной x; в самом деле, если эта функция при $x = n$ (целое) принимает некоторое значение N, то при всех значениях x, бóльших чем n (в том числе и дробных), она принимает значение, большее чем N.

---

[1]) В самом деле, достаточно обозначить x через c и сослаться на то, что функция x^c — возрастающая при $c > 0$ и убывающая при $c < 0$ (для случая c рационального — см. § 18, пример 4, для случая c иррационального — см. § 51).

Положив

$$a-1=\delta, \quad a=1+\delta \qquad (\delta>0),$$

мы получаем по формуле бинома

$$a^n=(1+\delta)^n=1+n\delta+\ldots+\delta^n>n\delta$$

(сумма положительных слагаемых больше одного второго слагаемого), и так как $n\delta$ при неограниченном возрастании n тоже неограниченно возрастает, то тем более такое заключение справедливо относительно a^n.

Пусть теперь p — какое-нибудь целое положительное число. Если мы хотим доказать, что при достаточно больших значениях n справедливо неравенство

$$a^n>n^p,$$

то достаточно применить прежнее рассуждение, выделяя, однако, вместо второго $(p+2)$-й член суммы:

$$a^n=(1+\delta)^n=1+n\delta+\ldots+\frac{n(n-1)\ldots(n-p)}{1\cdot 2\ldots(p+1)}\delta^{p+1}+\ldots+\delta^n>$$
$$>\frac{n(n-1)\ldots(n-p)}{1\cdot 2\ldots(p+1)}\delta^{p+1}. \qquad (72)$$

Но при достаточно больших значениях n

$$\frac{n(n-1)\ldots(n-p)}{1\cdot 2\ldots(p+1)}\delta^{p+1}>n^p. \qquad (73)$$

Действительно, это неравенство равносильно следующему:

$$n>\frac{1\cdot 2\ldots(p+1)}{\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{p}{n}\right)\delta^{p+1}},$$

а последнее очевидно, так как при неограниченном возрастании n правая часть приближается к числу $\frac{1\cdot 2\ldots(p+1)}{\delta^{p+1}}$, а левая неограниченно возрастает. Остаётся сопоставить неравенства (72) и (73). Например, полагая $\delta=0{,}01$, $p=100$, можно сказать, что при достаточно больших x

$$(1{,}01)^x>x^{100}.$$

Мы доказали, что *при неограниченном возрастании x показательная функция $y=a^x$* (при $a>1$) *тоже неограниченно возрастает, и притом быстрее, чем любая положительная степень x.*

Из тождества

$$a^{-x}=\left(\frac{1}{a}\right)^x$$

вытекает, что, *если x, оставаясь отрицательным, неограниченно возрастает по абсолютной величине, то показательная функция*

$y=a^x$ (при $a<1$) неограниченно убывает, и притом быстрее, чем любая отрицательная степень x.

Свойством показательной функции быстро возрастать или убывать при возрастании независимой переменной можно воспользоваться для доказательства её трансцендентности.

Доказательство. Предполагая для определённости, что $a>1$, допустим, что функция $y=a^x$ — алгебраическая. Это значит, что она тождественно удовлетворяет уравнению вида $P(x, y)=0$. Разложив многочлен $P(x, y)$ по убывающим степеням y, мы получим:

$$P(x, y)\equiv P_0(x)y^n+P_1(x)y^{n-1}+\ldots+P_{n-1}(x)y+P_n(x),$$

причём степень n — целая положительная и $P_0(x), P_1(x), \ldots, P_{n-1}(x), P_n(x)$ — многочлены относительно x; положим, в частности,

$$P_0(x)=Ax^m+\ldots,$$

где m — целое положительное число или нуль, но A отлично от нуля, и, не ограничивая общности, можно считать, что $A>0$.

Тождество

$$P(x, a^x)=0$$

можно переписать в виде

$$P_0(x)a^{nx}+P_1(x)a^{(n-1)x}+\ldots+P_{n-1}(x)a^x+P_n(x)=0,$$

или

$$P_0(x)=-P_1(x)a^{-x}-\ldots-P_{n-1}(x)a^{-(n-1)x}-P_n(x)a^{-nx}.$$

Предположим, что x неограниченно возрастает. Тогда многочлен $P_0(x)$ в левой части тождества также неограниченно возрастает (при $m>0$) или сводится к постоянной положительной величине A (при $m=0$). Что же касается правой части, то она распадается на сумму конечного числа членов, из которых каждый имеет вид Cx^sa^{-kx} $(k>0)$ и, следовательно, по доказанному свойству показательной функции, стремится к нулю (следует обратить внимание на то, что $a^{-k}=\frac{1}{a^k}<1$); поэтому и вся правая часть стремится к нулю. Получается противоречие, которое и доказывает утверждение о трансцендентности функции a^x.

Говоря о свойствах показательной функции, следует в особенности указать на её основное функциональное свойство, выражающееся в тождестве

$$a^{x'+x''}=a^{x'}a^{x''}, \tag{74}$$

справедливое при всех значениях x' и x'' («теорема сложения»).

Это соотношение очевидно для случая целых значений x' и x''. Оно доказывается для случая рациональных значений x' и x'', по приведении к общему знаменателю,

$$x'=\frac{p'}{q},\quad x''=\frac{p''}{q},$$

посредством использования свойств арифметических корней: из равенства

$$a^{p'+p''}=a^{p'}a^{p''}$$

следует равенство

$$a^{\frac{p'}{q}+\frac{p''}{q}}=\sqrt[q]{a^{p'+p''}}=\sqrt[q]{a^{p'}a^{p''}}=\sqrt[q]{a^{p'}}\sqrt[q]{a^{p''}}=a^{\frac{p'}{q}}a^{\frac{p''}{q}}.$$

Применяя теорему сложения для показательной функции в том случае, когда в показателе стоит сумма n слагаемых, из которых каждое равно x (с привлечением полной индукции), мы получаем тождество

$$a^{nx} = (a^x)^n. \tag{75}$$

Его легко обобщить: при любом (действительном) λ справедливо тождество

$$a^{\lambda x} = (a^x)^\lambda. \tag{76}$$

Для случая целого $\lambda\,(=n)$ это равенство уже доказано; если $\lambda = \frac{p}{q}$, то положим в соотношении (75) $n = p$ и извлечём арифметические корни степени q.

Равенство (75) говорит, между прочим, о следующем: *если независимое переменное x принимает ряд значений, образующих арифметическую прогрессию*

$$g,\ g+h,\ g+2h,\ \ldots,\ g+nh,\ \ldots,$$

то показательная функция a^x принимает ряд значений, образующих геометрическую прогрессию

$$G,\ Gq,\ Gq^2,\ \ldots,\ Gq^n,\ \ldots\ ^1).$$

Именно, мы получаем:

$$G = a^g,\ q = a^h.$$

Этому соответствует такое свойство графика функции a^x: ординаты, восставленные в равноотстоящих точках оси Ox, образуют геометрическую прогрессию.

Полагая в соотношении (74) $x' = x$, $x'' = -c$, мы получаем тождество

$$a^{x-c} = Ca^x,$$

(где $C = a^{-c}$). Это означает (см. § 3): *перемещение графика функции a^x параллельно оси Ox на отрезок с равносильно его растяжению по направлению оси Oy в a^{-c} раз* (сжатию в a^c раз).

В соотношении (76) как λ, так и x представляют собой совершенно произвольные числа: это даёт право поменять местами буквы λ и x:

$$a^{\lambda x} = (a^\lambda)^x.$$

Положив затем $\lambda = \frac{1}{p}$ $(p > 0)$, получим следующее свойство графика показательной функции a^x: *растяжение в p раз по направлению оси Ox равносильно переходу от графика показательной функции с основанием a к графику показательной функции с основанием $a^{\frac{1}{p}}$.*

---

[1]) Следует отметить, что, помимо показательных функций a^x, этим же свойством обладают и функции несколько более обширного класса ca^x (где c — постоянная).

Основному соотношению (74) можно также дать геометрическое истолкование. Именно, положив в нём $x' = \lambda x$, $x'' = \mu x$, мы получаем:

$$a^{\lambda x + \mu x} = a^{\lambda x} a^{\mu x}$$

и, далее, используя (76):

$$(a^{\lambda})^x (a^{\mu})^x = (a^{\lambda+\mu})^x, \quad \text{или} \quad a'^x a''^x = (a'a'')^x,$$

где

$$a' = a^{\lambda},$$
$$a'' = a^{\mu}.$$

Таким образом, *в результате «умножения» графиков показательных функций с основаниями a' и a''* (см. § 3) *возникает график показательной же функции с основанием, равным $a'a''$.*

§ 21. Функции, связанные с показательной

Рассмотрим функции

$$f(x) = \frac{a^x + a^{-x}}{2}, \qquad (a > 1).$$
$$g(x) = \frac{a^x - a^{-x}}{2}$$

Чтобы построить график функции $f(x)$, достаточно «сложить» (см. § 2, п. 1) взаимно симметричные относительно оси Oy графики показательных функций с основаниями a и $\frac{1}{a}$, затем произвести сжатие вдвое по направлению оси Oy. Получаемый график можно назвать «полусуммой» данных графиков. Особенно легко произвести построение по точкам, деля пополам точкой M отрезок M_1M_2, образованный на каждой вертикальной прямой графиками a^x и a^{-x}.

Чтобы построить график функции $g(x)$, нужно проделать то же самое, предварительно заменив график a^{-x} графиком $(-a^{-x})$, т. е. отразив его по симметрии относительно оси Ox.

Легко проверить, что:

1) функция $f(x)$ — чётная, функция $g(x)$ — нечётная; отсюда следуют свойства симметрии их графиков (см. § 3),

2) $f(0) = 1$, $g(0) = 0$,

3) $f(x) > 0$ при всех значениях x,

$$g(x) \begin{cases} > 0 & \text{при } x > 0 \\ < 0 & \text{при } x < 0. \end{cases}$$

В самом деле, из $a > 1$ следует $a > a^{-1}$, и тогда $a^x > a^{-x}$ (см. § 20) при $x > 0$ и т. д.

4) При всех значениях x

$$[f(x)]^2 - [g(x)]^2 = 1. \quad (77)$$

Докажем, что

5) функция $f(x)$ — возрастающая при $x > 0$ и (по симметрии) убывающая при $x < 0$; функция $g(x)$ — всегда возрастающая.

Для доказательства достаточно заметить, что

$$f(x) = \frac{1}{2}\left(u + \frac{1}{u}\right), \quad g(x) = \frac{1}{2}\left(u - \frac{1}{u}\right),$$

где положено

$$u = a^x.$$

При $x > 0$ мы имеем $u > 1$, а при условии $u > 1$ функции $u + \frac{1}{u}$ и $u - \frac{1}{u}$ являются возрастающими (см. § 15, (I) и (II)). Так как показательная функция $u = a^x$ при $a > 1$ — также возрастающая, то отсюда следует (§ 5, п. 9) требуемое заключение.

Опираясь на основное функциональное свойство показательной функции, можно вывести соответствующие свойства функций $f(x)$ и $g(x)$.

Именно,

$$f(x' + x'') = \frac{1}{2}(a^{x'+x''} + a^{-(x'+x'')}) = \frac{1}{2}(a^{x'}a^{x''} + a^{-x'}a^{-x''}),$$

и с другой стороны —

$$f(x')f(x'') + g(x')g(x'') =$$

$$= \frac{a^{x'} + a^{-x'}}{2} \cdot \frac{a^{x''} + a^{-x''}}{2} + \frac{a^{x'} - a^{-x'}}{2} \cdot \frac{a^{x''} - a^{-x''}}{2} = \frac{1}{2}(a^{x'}a^{x''} + a^{-x'}a^{-x''});$$

значит,

$$f(x' + x'') = f(x')f(x'') + g(x')g(x''). \quad (78)$$

Аналогично доказывается, что

$$g(x' + x'') = f(x')g(x'') + g(x')f(x''). \quad (79)$$

Рассмотрим ещё функцию

$$h(x) = \frac{g(x)}{f(x)} = \frac{a^x - a^{-x}}{a^x + a^{-x}} \qquad (a > 1). \quad (80)$$

Так как функция $f(x)$ — чётная, а функция $g(x)$ — нечётная, то функция $h(x)$ — нечётная, причём $h(0) = 0$. Очевидно, $h(x) > 0$ при $x > 0$ и $h(x) < 0$ при $x < 0$.

Функция $h(x)$ при $x > 0$ — возрастающая. В самом деле, если x возрастает, начиная от значения нуль, неограниченно, то $u = a^{-2x}$ убывает от 1, неограниченно приближаясь к нулю, и потому (опи-

раясь на § 5, п. 6, 1, 7 и 4), из записи

$$h(x)=\frac{1-u}{1+u}$$

мы видим, что при неограниченном возрастании x функция $h(x)$ возрастает и притом приближается к 1. Из свойства симметрии видно, что $h(x)$ возрастает на всей оси и что, если x неограниченно возрастает по абсолютной величине, оставаясь отрицательным, то $h(x)$ приближается к -1.

Для функций $f(x)$, $g(x)$ и $h(x)$ приняты (обычно — только при $a=e=2{,}718\ldots$, см. § 44) обозначения и наименования:

$f(x)=\operatorname{ch} x$ («косинус гиперболический») [1]),

$g(x)=\operatorname{sh} x$ («синус гиперболический»),

$h(x)=\operatorname{th} x$ («тангенс гиперболический»).

В этих обозначениях тождества (77) — (80) принимают вид

$$\operatorname{ch}^2 x-\operatorname{sh}^2 x=1, \tag{81}$$

$$\operatorname{ch}(x'+x'')=\operatorname{ch} x'\operatorname{ch} x''+\operatorname{sh} x'\operatorname{sh} x'', \tag{82}$$

$$\operatorname{sh}(x'+x'')=\operatorname{ch} x'\operatorname{sh} x''+\operatorname{sh} x'\operatorname{ch} x'' \text{ [2])}, \tag{83}$$

$$\operatorname{th} x=\frac{\operatorname{sh} x}{\operatorname{ch} x}. \tag{84}$$

Здесь ясно видна формальная аналогия с обыкновенными («круговыми») функциями $\cos x$, $\sin x$ и $\operatorname{tg} x$; отсюда происходят наименования «косинус», «синус», «тангенс» [3]).

Что касается термина «гиперболические» функции, то он объясняется следующим образом. Как известно из тригонометрии (см. также § 25), функции

$$X=\cos x,\quad Y=\sin x$$

удовлетворяют тождественно уравнению

$$X^2+Y^2=1.$$

Последнее уравнение (в плоскости OXY) представляет окружность, и потому эти функции называются «круговыми». Совершенно

---

[1]) Для краткости будем в этом параграфе функции $f(x)$, $g(x)$ и $h(x)$ называть «гиперболическим косинусом», «гиперболическим синусом» и «гиперболическим тангенсом» при любом a ($a>0$).

[2]) Тождества (82) и (83) называются «теоремой сложения» (для гиперболических функций $\operatorname{ch} x$ и $\operatorname{sh} x$).

[3]) Причина возникновения формальных аналогий между круговыми и гиперболическими функциями объяснена на стр. 507—508.

таким же образом функции

$$X = \operatorname{ch} x \quad \text{и} \quad Y = \operatorname{sh} x$$

удовлетворяют уравнению

$$X^2 - Y^2 = 1$$

(см. формулу (81)). Так как это уравнение в плоскости OXY пред-

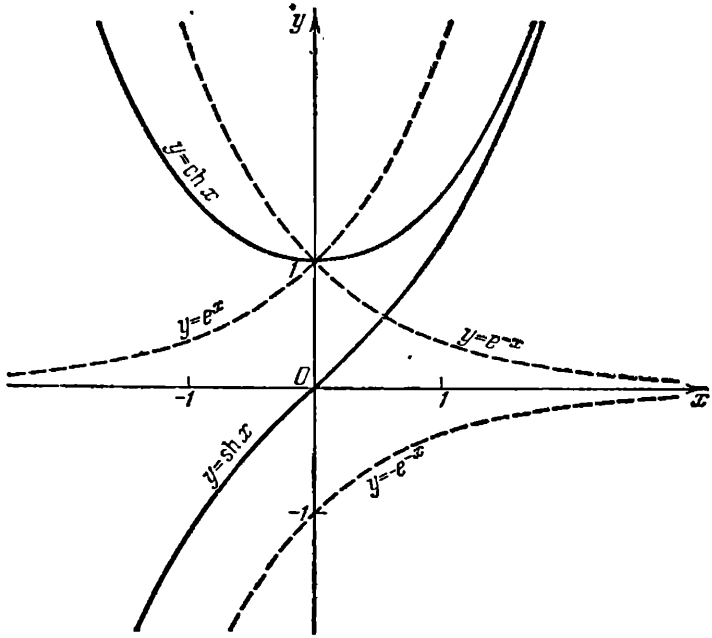


Рис. 38.

ставляет гиперболу (§ 18, пример 9), то отсюда происходит термин «гиперболические» функции.

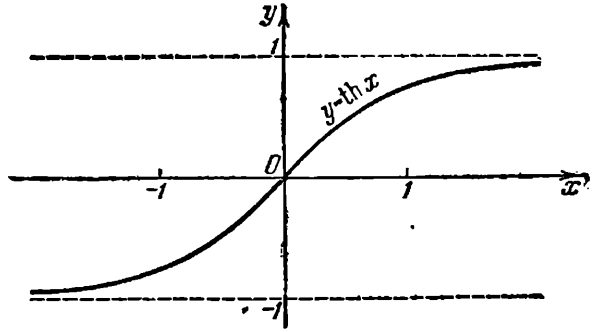


Рис. 39.

Графики функций $\operatorname{sh} x$ и $\operatorname{ch} x$ изображены на рис. 38; график функции $\operatorname{th} x$ — на рис. 39.

§ 22. Логарифмическая функция

Логарифмическая функция (короче, *логарифм*) *по основанию* a ($a > 0$) определяется как функция, обратная показательной, с тем же основанием.

Допустим, что $a > 1$. Если показательная функция задаётся уравнением

$$y = a^x \qquad (a > 1), \tag{85}$$

то уравнение, определяющее логарифмическую функцию, получается из него посредством перестановки букв x и y, что соответствует изменению роли переменных (см. § 4). Логарифмическая функция по основанию a

$$y = \log_a x \tag{86}$$

определяется из уравнения $x = a^y$.

График уравнения (86) симметричен графику уравнения (85) относительно биссектрисы $y = x$ (см. § 4). На рис. 40 изображены:

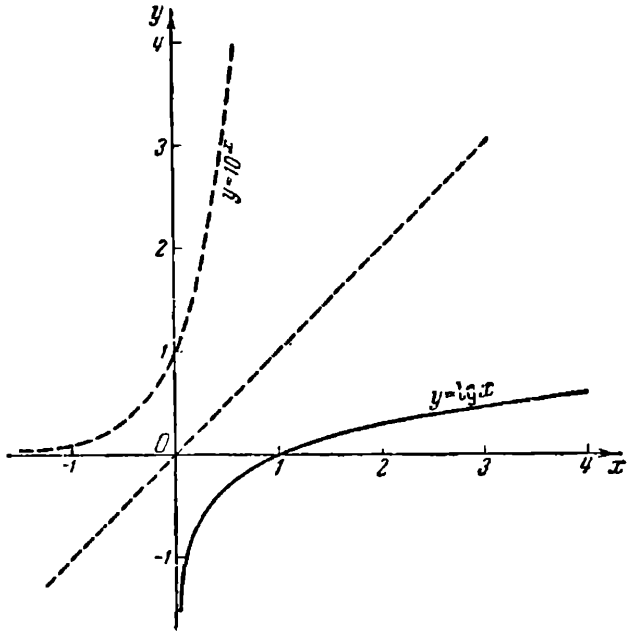


Рис. 40.

график функции $y = \lg x$ (сплошная линия) и симметричный ему относительно биссектрисы $y = x$ график функции $y = 10^x$ (пунктирная линия).

Принципиальный вопрос о том, соответствует ли в силу уравнения (85) заданному значению x одно определённое значение y, кажется допускающим очевидный ответ, именно утвердительный, в случае, если x положительно; ответ отрицательный в случае, если x отрицательно или равно нулю. В самом деле, речь идёт о том,

пересекается ли в единственной точке с графиком (85) данная вертикальная прямая, и чертёж подсказывает решение. Логическое обоснование существования логарифма положительного числа по данному положительному основанию приведено ниже (см. § 52).

Все свойства обратной функции — логарифма непосредственно вытекают из свойств прямой функции — показательной, и иллюстрируются рис. 40.

Прежде всего, как уже было только что отмечено, *существуют логарифмы* $\log_a x$ *только тех чисел* x, *которые положительны* ($x > 0$) (так как показательная функция принимает лишь положительные значения).

Функция $\log_a x$ при $x > 0$ оказывается, далее, возрастающей (см. § 52). Она возрастает, с увеличением x, неограниченно, однако медленнее, чем любая положительная степень x. Действительно, пусть δ — постоянное положительное, сколь угодно малое число. Тогда, полагая

$$\log_a x = y, \quad x = a^y,$$

мы получим:

$$\frac{\log_a x}{x^\delta} = \frac{y}{a^{\delta y}} = \left(\frac{y^{\frac{1}{\delta}}}{a^y}\right)^\delta.$$

По свойству показательной функции, при достаточно больших значениях y величина a^y становится больше, чем любая степень y, например $y^{\frac{1}{\delta}}$. Значит, выражение в скобках становится меньше единицы, откуда следует (при достаточно больших значениях x), что

$$\log_a x < x^\delta.$$

Когда же x приближается к нулю, то функция $\log_a x$, делаясь (при $x < 1$) отрицательной, по абсолютному значению неограниченно возрастает, однако медленнее, чем любая отрицательная степень x.

Нам нет необходимости задерживаться на замечательном функциональном свойстве функции логарифм

$$\log_a (x'x'') = \log_a x' + \log_a x'', \tag{87}$$

вытекающем из функционального свойства («теорема сложения») показательной функции (74) и лежащем в основе вычислительных применений логарифмов. Излишне также перечислять достаточно известные следствия, вытекающие из формулы (87).

Отметим ещё некоторые свойства графиков логарифмических функций. Из формулы

$$\log_a \frac{x}{p} = \log_a x - \log_a p \tag{88}$$

следует: *растяжение графика функции* $\log_a x$ *в* p *раз по направлению оси* Ox *равносильно перенесению его параллельно оси* Oy *на отрезок* $(-\log_a p)$.

Формула

$$q \log_a x = \log_{a^{1/q}} x \tag{89}$$

говорит о том, что *растяжение графика логарифма по основанию a в q раз по направлению оси Oy равносильно переходу от этого графика к графику логарифма по основанию $a^{\frac{1}{q}}$.*

Логарифмируя тождества

$$a^{\log_a x} = x, \quad b^{\log_b x} = x$$

соответственно по основаниям b и a, мы получаем

$$\log_a x \cdot \log_b a = \log_b x, \quad \log_b x \cdot \log_a b = \log_a x,$$

или

$$\log_a x = \frac{\log_b x}{\log_b a}, \quad \log_b x = \frac{\log_a x}{\log_a b}. \tag{90}$$

Эти формулы показывают, как логарифм числа по некоторому основанию выражается через логарифм того же числа по другому основанию; из них, между прочим, следует (если положим $x = b$ в первой из формул или $x = a$ во второй), что $\log_a b$ и $\log_b a$ — величины, взаимно обратные:

$$\log_a b \cdot \log_b a = 1.$$

Таким образом, *чтобы перейти от системы логарифмов по одному основанию к системе по другому основанию, достаточно умножить логарифмы на некоторый постоянный множитель* («*модуль перехода*»). Этому как раз соответствует геометрически растяжение графика логарифма по направлению оси Oy (нужно положить в формуле (89) $a^{\frac{1}{q}} = b$).

Примечание. Мы предположили, что $a > 1$. Если $0 < a < 1$, то вопрос о логарифме по основанию a исчерпывается тем, что имеет место тождество

$$\log_a x = -\log_{\frac{1}{a}} x.$$

(В самом деле, если $x = a^y$, то $x = \left(\frac{1}{a}\right)^{-y}$.) Случай же $a = 1$, очевидно, не заслуживает рассмотрения.

§ 23. Функции, связанные с логарифмической

1. Поскольку функции $f(x)$, $g(x)$ и $h(x)$, введённые в § 21, являются простыми комбинациями из показательной, неудивительно, что функции, им обратные, оказываются связанными с логарифмической.

Найдём явное выражение для функций, обратных гиперболическим.

Нам придётся (§ 4) уравнениях

$$y=\frac{a^x+a^{-x}}{2},\quad y=\frac{a^x-a^{-x}}{2}, \tag{91}$$

$$y=\frac{a^x-a^{-x}}{a^x+a^{-x}} \tag{92}$$

поменять местами буквы x и y и затем решить полученные уравнения

$$x=\frac{a^y+a^{-y}}{2},\quad x=\frac{a^y-a^{-y}}{2}, \tag{93}$$

$$x=\frac{a^y-a^{-y}}{a^y+a^{-y}} \tag{94}$$

относительно y.

Станем решать уравнения (93) параллельно. После умножения на $2a^y$ эти уравнения принимают вид

$$a^{2y}-2xa^y+1=0,\quad a^{2y}-2xa^y-1=0,$$

т. е. представляют собой квадратные уравнения относительно a^y, так что дальше отсюда следует:

$$a^y=x\pm\sqrt{x^2-1},\quad a^y=x+\sqrt{x^2+1}.$$

Во втором случае знак минус перед радикалом излишен, так как, решая уравнение относительно a^y, мы, естественно, разыскиваем только положительные его корни; выражение же $x-\sqrt{x^2+1}$ заведомо отрицательно (относительно разности $x-\sqrt{x^2-1}$ этого сказать нельзя). Дальше остаётся прологарифмировать:

$$y=\log_a(x\pm\sqrt{x^2-1}),\quad y=\log_a(x+\sqrt{x^2+1}).$$

Замечая, что

$$x-\sqrt{x^2-1}=\frac{1}{x+\sqrt{x^2-1}},$$

получаем функции, обратные «гиперболическим косинусу и синусу» (91), в виде

$$y=\pm\log_a(x+\sqrt{x^2-1}),\quad y=\log_a(x+\sqrt{x^2+1}). \tag{95}$$

Первая из них задана при ограничении $x\geqslant 1$ и двузначна (два значения различаются знаками); вторая задана без ограничений и однозначна.

Что касается уравнения (94), то, решая его относительно a^y, получаем:

$$a^y=\sqrt{\frac{1+x}{1-x}}\ \text{(радикал — арифметический)},$$

откуда видно, что функция, обратная «гиперболическому тангенсу» (92),

есть

$$y=\frac{1}{2}\log_a\frac{1+x}{1-x}. \tag{96}$$

Она задаётся с ограничением $-1<x<+1$ и однозначна.

В частности, при $a=e$ формулы (95) и (96) дают нам функции, обратные гиперболическим косинусу, синусу и тангенсу (в собственном смысле):

$$\operatorname{arch} x=\pm\log_e\left(x+\sqrt{x^2-1}\right),\quad \operatorname{arsh} x=\log_e\left(x+\sqrt{x^2+1}\right), \tag{95'}$$

$$\operatorname{arth} x=\frac{1}{2}\log_e\frac{1+x}{1-x}. \tag{96'}$$

Графики функций (95) и (96) (для случая $a=e$) получаются из графиков функций (91) и (92) (см. рис. 38 и 39) посредством симметричного отражения относительно биссектрисы $y=x$.

2. Часто бывает нужно отдавать себе отчёт в поведении логарифма данной функции (по основанию a), зная поведение самой функции; другими словами, по графику данной функции наметить график её логарифма.

Это сделать нетрудно, если не упускать из виду следующих, достаточно очевидных, обстоятельств [1]):

1) функция $\log_a f(x)$ существует при условии (и только при том условии), что функция $f(x)$ существует и положительна;

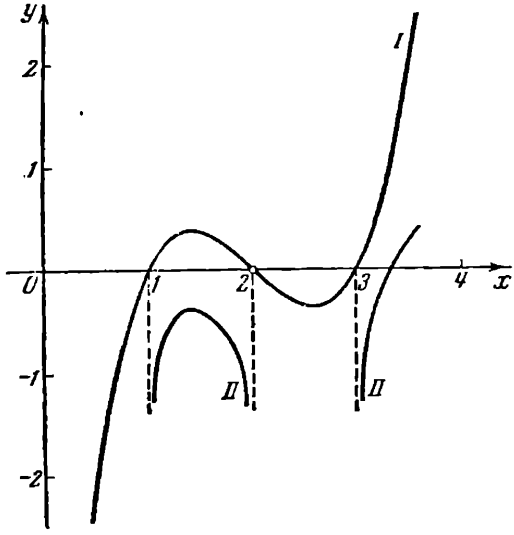


Рис. 41.

2) функция $\log_a f(x)$ равна нулю, положительна или отрицательна, смотря по тому, будет ли функция $f(x)$ равна единице, больше или меньше единицы; вообще же при $a=10$ (а также и при $a=e$)

$$\log_a f(x)<f(x);$$

3) функция $\log_a f(x)$ возрастает (или убывает) в тех же промежутках, что и функция $f(x)$ (см. § 5, п. 9); она имеет максимум (или минимум) в тех же точках, что и $f(x)$.

Последнее отмеченное обстоятельство особенно важно в том отношении, что нередко бывает гораздо легче найти максимум или минимум логарифма функции, чем самой функции.

На рис. 41 изображены совместно графики:

а) функции $f(x)=(x-1)(x-2)(x-3)$ (кривая *I*),

б) функции $\log_{10} f(x)=\log_{10}[(x-1)(x-2)(x-3)]$ (кривая *II*).

[1]) Попрежнему предполагаем дальше, что $a>1$.

§ 24. Произвольная степенная функция

Степенная функция

$$y = x^{\alpha} \tag{97}$$

была рассмотрена для случаев, когда α — целое положительное или отрицательное число (см. § 7 и § 13), затем — для случая, когда $\alpha = \frac{p}{q}$ — любое рациональное число (§ 18). Случай когда α — иррациональное число, подразумевает предельный переход «по непрерывности» в показателе. Заметим предварительно, что функция x^{α} даже в случае α рационального (при чётном знаменателе) теряет смысл для отрицательных значений x; если α — иррациональное, то при x отрицательном выражению x^{α} нельзя приписать никакого смысла ни непосредственно, ни в результате предельного перехода. По этой причине, рассматривая иррациональную степень x^{α}, переменной x не дают отрицательных значений и считают, что эта функция задана только для положительных значений x. Можно доказать, что функция (97) при α иррациональном является трансцендентной функцией, т. е. степенная функция относится к числу алгебраических функций только при α рациональном.

Выбрав произвольное положительное основание a, формуле (97) часто придают вид

$$x^{\alpha} = (a^{\log_a x})^{\alpha} = a^{\alpha \log_a x}\ ^{1}). \tag{98}$$

Такая запись имеет теоретическое и практическое оправдание. С одной стороны, предельный переход по непрерывности представляет собой довольно сложное построение, которое можно осуществить одинаково как по отношению к показательной функции, так и по отношению к степенной (в последнем случае, как было указано, с ограничением $x > 0$); но целесообразно осуществить её лишь один раз, именно, по отношению к показательной функции, с дальнейшим автоматическим перенесением на обратную функцию — логарифмическую (см. § 52) и, далее, опять автоматически — с помощью формулы (98) — на степенную [2]).

---

[1]) Логически это — определение произвольной степени; формула логарифмирования произвольной степени отсюда вытекает как следствие.

[2]) Возможен и иной ход мыслей: сначала определяется для всех значений $x > 1$ логарифм («натуральный», т. е. по основанию e, см. гл. III, § 44) согласно формуле

$$\ln x = \int_1^x \frac{d\xi}{\xi}$$

как площадь, ограниченная гиперболой $y = \frac{1}{x}$, осью Ox и вертикалями, проведёнными через точки 1 и x на этой оси; затем показательная функция (с основанием e) определяется как обратная по отношению к логарифму; наконец, степенная функция x^{α} определяется по формуле $x^{\alpha} = e^{\alpha \ln x}$.

С другой стороны, практическое вычисление значений функции x^α при ряде данных значений x удобнее всего произвести

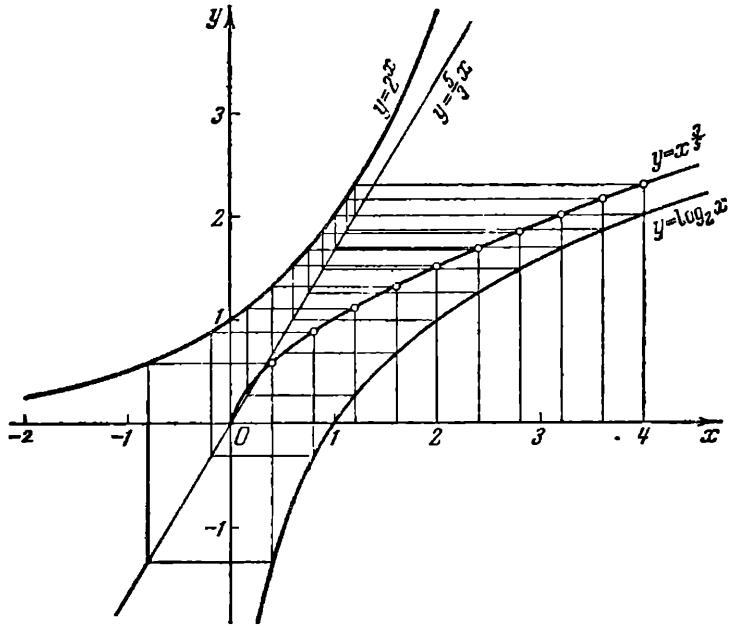


Рис. 42.

с помощью таблицы логарифмов, составленной по некоторому основанию a, а это и приводит по сути дела к формуле (98).

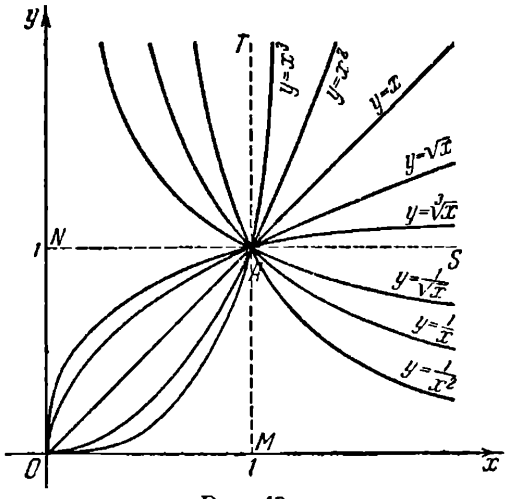


Рис. 43.

Основываясь на формуле (98), можно с помощью двух кривых — графиков функций a^x и $\log_a x$ осуществить точечное построение любой степенной функции x^α. Это построение ясно из рис. 42, на котором взято $a=2$ и выбрано значение $\alpha=\frac{3}{5}$.

По поводу графиков функций $y=x^\alpha$ (при произвольных значениях показателя α) полезно сделать следующие замечания (рис. 43):

1) Все они проходят через точку A (1, 1).

2) Если $\alpha>0$, то функция x^α возрастающая; график её, выходя из начала O и уходя в бесконечность, целиком расположен в квадрате $OMAN$ и квадранте SAT (см. рис. 43).

Если $\alpha<0$, то функция x^α — убывающая; график её целиком расположен в частях плоскости $xMAS$ и $yNAT$, асимптотически приближаясь к осям Ox и Oy.

3) Если $\alpha'<\alpha''$, то из двух кривых $y=x^{\alpha'}$ и $y=x^{\alpha''}$ первая лежит выше второй при $0<x<1$ и ниже второй при $x>1$. Это — следствие из свойств показательной функции.

4) Графики кривых $y=x^\alpha$ и $y=x^{\frac{1}{\alpha}}$ взаимно симметричны относительно биссектрисы $x=y$ (например, x^2 и $\sqrt{x}$; $\frac{1}{x^2}$ и $\frac{1}{\sqrt{x}}$).

§ 25. Основные (целые) тригонометрические функции: синус и косинус

Рассмотрим в плоскости OXY «единичный круг»

$$X^2+Y^2=1, \tag{99}$$

с центром O и радиусом 1.

Точку A пересечения окружности с положительной полуосью OX будем считать «начальной»; вообразим точку M, движущуюся (вращающуюся) по окружности, причём положительным направлением вращения будем считать направление от оси OX к оси OY, т. е. против часовой стрелки (рис. 44).

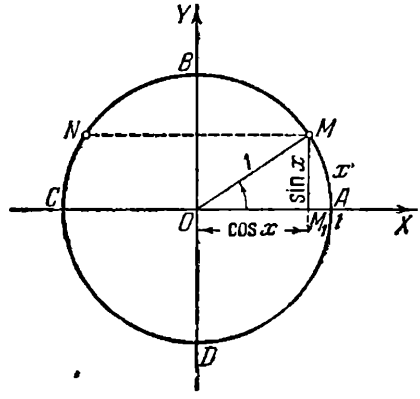


Рис. 44.

Положение точки M определяется однозначно, если указать длину дуги AM

$$\smile AM=x,$$

отсчитываемую в положительном направлении от начальной точки A до точки M. Когда x возрастает от 0 до 2π, точка M делает полный оборот. Промежутки изменения $0<x<\frac{\pi}{2}$, $\frac{\pi}{2}<x<\pi$, $\pi<x<\frac{3}{2}\pi$, $\frac{3}{2}\pi<x<2\pi$ (которым соответствуют дуги AB, BC, CD, DA; см. рис. 44) носят названия *первой, второй, третьей, четвёртой* четверти [1]).

[1]) Следующая, пятая, четверть $\left(2\pi<x<\frac{5}{2}\pi\right)$, которой снова, как и первой, соответствует дуга AB, «гомологична» первой; точно так же шестая четверть $\left(\frac{5}{2}\pi<x<3\pi\right)$ гомологична второй и т. д. «Минус первая» четверть $\left(-\frac{\pi}{2}<x<0\right)$ гомологична четвёртой, «минус вторая» $\left(-\pi<x<-\frac{\pi}{2}\right)$ третьей, и т. д. «Нулевой» четверти нет вовсе.

Абсцисса X и ордината Y точки M, расположенной на окружности таким образом, что длина дуги AM равна x, являются функциями величины x, называемыми *косинусом* и *синусом*:

$$X = \cos x = OM_1, \quad Y = \sin x = M_1M. \tag{100}$$

Из соотношения (99) следует тождество

$$\cos^2 x + \sin^2 x = 1. \tag{101}$$

Таким образом, определение основных тригонометрических функций, сообщаемое в школе, носит геометрический характер. Ввиду того, что оно связано с рассмотрением окружности (круга), тригонометрические функции иначе называются *круговыми*. Существуют и различные аналитические (формульные) определения синуса и косинуса: простейшие из формул, которые могут быть взяты в качестве определений, — степенные ряды (см. стр. 471 и 500) подразумевают выполнение лишь двух действий — сложения и умножения, но число этих действий бесконечно (что равносильно наличию предельного перехода, см. § 48); как известно, о подобного рода формулах школьные программы не упоминают.

Замечательное свойство синуса и косинуса, которое отличает их от всех ранее рассмотренных нами функций, — их периодичность. Геометрически ясно, что после полного оборота по окружности точка M снова оказывается на прежнем месте; отсюда следуют тождества

$$\sin(x + 2\pi) = \sin x, \quad \cos(x + 2\pi) = \cos x, \tag{102}$$

свидетельствующие о том (см. § 3), что 2π есть период функций синуса и косинуса [1]).

Наличие периода позволяет сделать заключение о трансцендентности тригонометрических функций (см. § 19). В самом деле, раз функция имеет период, то она принимает одно и то же значение c в бесконечном ряде различных точек: например, $\cos x$ принимает значение 1 в точках вида $2k\pi$ [2]). Но функция, обладающая этим свойством и не сводящаяся к постоянной c, никак не может быть алгебраической. Действительно, алгебраическая функция $y = f(x)$ определяется уравнением вида

$$P(x, y) = 0,$$

где $P(x, y)$ — многочлен относительно x и y; допустим, что y имеет значение c при бесчисленном множестве значений x, но не при всех значениях x; тогда алгебраическое уравнение

$$P(x, c) = 0, \tag{35'}$$

---

[1]) Итак, в «гомологических» точках (различающихся на величину, кратную 2π) каждая из функций $\cos x$ и $\sin x$ принимает одни и те же значения. При таких условиях во многих случаях можно, не различая гомологических четвертей, ограничиться рассмотрением первых четырёх, образующих один период.

[2]) Здесь и дальше k обозначает произвольное целое число.

не обращаясь в тождество относительно x, имеет бесчисленное множество корней. Это, однако, невозможно, так как уравнение (35') относительно x — алгебраическое и степень его, очевидно, не превышает степени уравнения (35) относительно пары переменных x и y.

Итак, функции синус и косинус — трансцендентные.

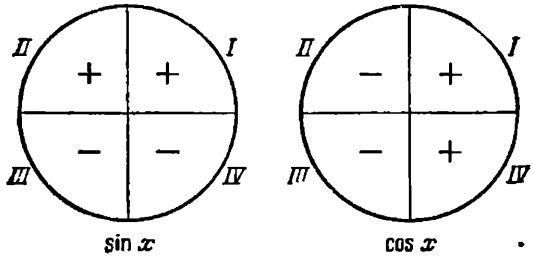
О знаке основных тригонометрических функций, а также об их возрастании или убывании следует судить, исходя из определения, т. е. основываясь на геометрических представлениях. Функция $\sin x$ обращается в нуль тогда и только тогда, когда равна нулю ордината точки M, т. е. сама точка оказывается лежащей на оси Ox; итак,

$$\sin k\pi = 0. \tag{103}$$

Функция $\cos x$ обращается в нуль тогда и только тогда, когда равна нулю абсцисса точки M, т. е. сама точка оказывается лежащей на оси Oy; итак,

$$\cos\left(\frac{\pi}{2} + k\pi\right) = 0. \tag{104}$$

Знаки $\sin x$ и $\cos x$ (по четвертям) определяются в зависимости от знаков ординаты и абсциссы точки M, а именно, согласно схемам:



Обе функции синус и косинус способны изменяться лишь в пределах от -1 до $+1$. При этом синус принимает наибольшее значение $+1$ на границе первой и второй четверти

$$\sin\left(\frac{\pi}{2} + 2k\pi\right) = 1, \tag{105}$$

а наименьшее значение -1 — на границе третьей и четвёртой

$$\sin\left(\frac{3}{2}\pi + 2k\pi\right) = -1. \tag{106}$$

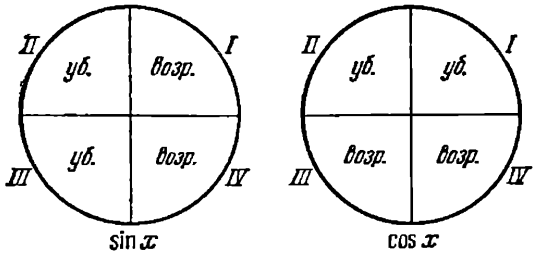
Что же касается косинуса, то он принимает наибольшее значение $+1$ на границе четвёртой и первой четверти

$$\cos 2k\pi = 1, \tag{107}$$

а наименьшее — на границе второй и третьей

$$\cos(\pi + 2k\pi) = -1. \tag{108}$$

Изменяемость синуса и косинуса по четвертям даётся схемами:



Наилучший способ для запоминания указанных свойств — непосредственно опираться на зрительные представления единичного круга с вращающейся точкой (см. рис. 44). Этот простой рисунок всегда легко воспроизвести на бумаге или мысленно.

Не излишне, впрочем, сопоставляя обе схемы, обратить внимание на то, что изменяемость синуса находится в «прямом соответствии» со знаком косинуса (т. е. синус возрастает там, где косинус положителен, и убывает там, где косинус отрицателен), а изменяемость косинуса — в «обратном соответствии» со знаком синуса [1]).

Отмерим на единичном круге (рис. 45) от точки A дуги x и $-x$ и в концах их поставим точки M и M_1; эти точки имеют одну и ту же абсциссу, но их ординаты отличаются знаком. Отсюда следует (см. § 3), что косинус — чётная функция, а синус — нечётная:

$$\cos(-x)=\cos x, \quad \sin(-x)=-\sin x. \tag{109}$$

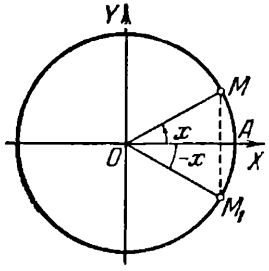


Рис. 45.

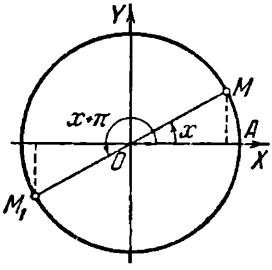


Рис. 46.

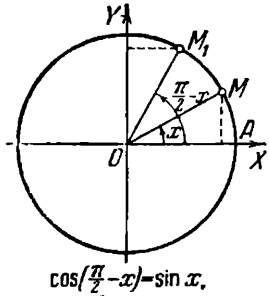


Рис. 47.

Рассмотрим дальше точки M и M_1, находящиеся в концах дуг x и $x+\pi$ (рис. 46). Эти точки расположены на противоположных концах диаметра единичного круга и, значит, симметричны относи-

[1]) Это — лишь часть того, что содержится в «правилах дифференцирования»:

$$(\sin x)'=\cos x, \quad (\cos x)'=-\sin x$$

(см. стр. 312, формулы 5) и 6)).

тельно центра O. Поэтому отличаются между собой лишь знаком их абсциссы, а также их ординаты. Следовательно,

$$\cos(x+\pi)=-\cos x, \quad \sin(x+\pi)=-\sin x. \tag{110}$$

Наконец, возьмём две точки M и M_1, находящиеся в концах дуг x и $\frac{\pi}{2}-x$ (рис. 47). Легко понять, что такие точки взаимно симметричны относительно биссектрисы $x=y$. Значит (см. § 4), абсцисса M_1 равна ординате M, а ордината M_1 — абсциссе M. Итак,

$$\cos\left(\frac{\pi}{2}-x\right)=\sin x, \quad \sin\left(\frac{\pi}{2}-x\right)=\cos x. \tag{111}$$

Из формул (111), (109), (110) легко получается следующая:

$$\cos x=\sin\left(x+\frac{\pi}{2}\right). \tag{112}$$

Переходя к построению и исследованию графиков функций $\sin x$ и $\cos x$, заметим, что достаточно получить график синуса в промежутке первой четверти $\left(0<x<\frac{\pi}{2}\right)$ для того, чтобы затем, пользуясь формулами (109—110), посредством элементарных преобразований продолжить его на всю ось $(-\infty<x<+\infty)$.

Действительно, по формуле (109) начало O есть центр симметрии этого графика, что позволяет продолжить его на промежуток $\left(-\frac{\pi}{2}<x<\frac{\pi}{2}\right)$; вторая из формул (110), говорящая о том, что прямая $x=\frac{\pi}{2}$ есть ось симметрии графика [1]), позволяет продолжить его на промежуток $\left(-\frac{\pi}{2}<x<\frac{3}{2}\pi\right)$; дальнейшее продолжение становится возможным вследствие существования периода 2π (формула (102)). Что касается графика косинуса, то, как видно из формулы (112), он получается из графика синуса посредством перенесения параллельно оси Ox на отрезок $\left(-\frac{\pi}{2}\right)$.

Чтобы выполнить точечное построение графика функции $\sin x$ в пределах первой четверти $\left(0<x<\frac{\pi}{2}\right)$, прибегают обыкновенно к делению на m равных промежутков.

---

[1]) Это заключение следует не непосредственно: заменяя в формуле (110) x на $x-\frac{\pi}{2}$ и затем используя нечётность синуса (формула (109)), мы получаем:

$$\sin\left(\frac{\pi}{2}+x\right)=\sin\left(\frac{\pi}{2}-x\right),$$

и последнее уже показывает, что прямая $x=\frac{\pi}{2}$ есть ось симметрии.

Разделим, с одной стороны, на m равных частей четверть дуги AB единичного круга (в плоскости OXY, рис.48, *а*); с другой, разделим на m же равных частей отрезок $\left(0, \frac{\pi}{2}\right)$ на оси Ox (в плоскости Oxy, рис. 48, *б*) и проведём через точки деления вертикали. Откладывая затем ординаты точек деления на дуге AB на соответствующих вертикалях в плоскости Oxy, получим ряд точек графика функции $\sin x$.

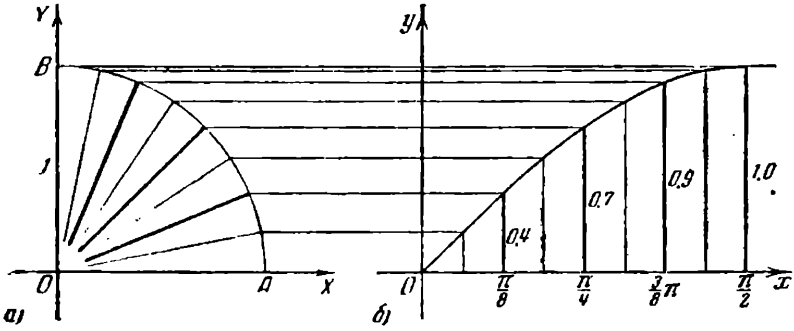


Рис. 48.

Если нужно наметить график синуса довольно быстро (и не особенно заботясь о точности), то очень полезно бывает построение «через четыре точки», соответствующее случаю $m = 4$. При этом длины ординат (значения $\sin x$ при $x = 0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3}{8}\pi, \frac{\pi}{2}$) оказываются приблизительно равными 0, 4, 7, 9, 10 десятых [1]). Эти числа заслуживают того, чтобы их запомнить (см. более жирные ординаты на рис. 48, *б*).

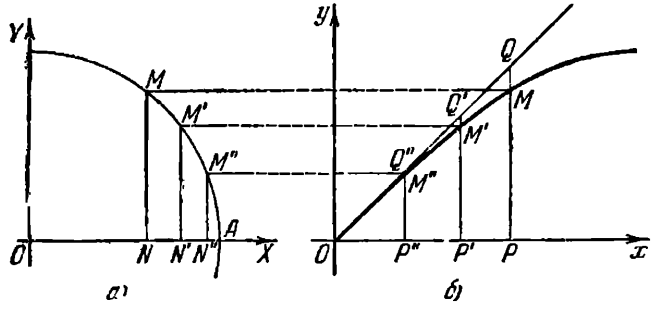


Рис. 49.

Существенно обратить внимание ещё на одно обстоятельство. Рассмотрим на единичном круге несколько положений M, M', M'' точки, движущейся к точке A, соответствующих весьма малым значениям независимой переменной x. Как видно на рис. 49, *а*, орди-

[1]) Точнее: 0; 0,383; 0,707; 0,924; 1.

ната NM точки M, конечно, несколько меньше, чем длина дуги AM. Но отношение $\frac{NM}{\smile AM}$ при приближении M к A неограниченно приближается к единице. Этому соответствует (рис. 49, *б*) тот факт, что ордината PM точки M на графике синуса, движущейся к началу O, хотя и несколько меньше, чем абсцисса OP (или, что то же, ордината PQ точки Q на биссектрисе $y=x$), однако с приближением M к O отношение $\frac{PM}{OP}$ $\left(\text{или } \frac{PM}{PQ}\right)$ всё меньше отличается от 1.

Таким образом, *график синуса лежит ниже биссектрисы* $y=x$:

$$\sin x < x \quad (x > 0);$$

но по мере приближения к началу O всё теснее к ней примыкает («касается»).

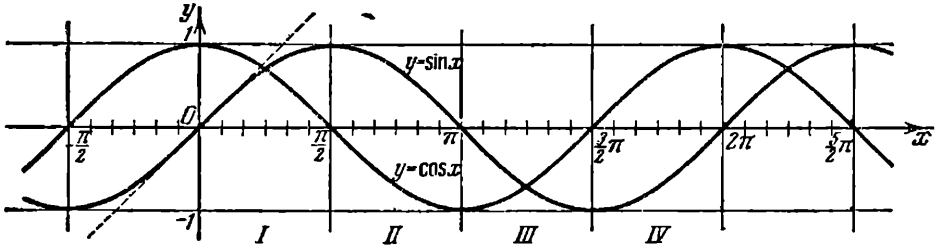


Рис. 50.

На рис. 50 показаны графики функций

$$y=\sin x \quad \text{и} \quad y=\cos x. \tag{113}$$

Эти кривые носят название *синусоиды* и *косинусоиды*.

«Теоремы сложения» для функций $\sin x$ и $\cos x$

$$\sin(x'+x'')=\sin x' \cos x''+\cos x' \sin x'',$$
$$\cos(x'+x'')=\cos x' \cos x''-\sin x' \sin x''$$

хорошо известны из курса тригонометрии; нам незачем на них останавливаться (см. также стр. 505).

§ 26. Простые гармонические колебания

График уравнения

$$y=C\sin\lambda(x-c) \quad (C>0,\ \lambda>0) \tag{114}$$

получается из графика $y=\sin x$ посредством следующих последовательно выполненных преобразований:

1) сжатия в λ раз (растяжения в $\frac{1}{\lambda}$ раз) по направлению оси Ox,
2) растяжения в C раз по направлению оси Oy,
3) перенесения параллельно оси Ox на отрезок, равный c.

Всякую функцию вида (114), а также её график, называют *простым гармоническим колебанием* (или просто *гармоническим колебанием*). Соответствующую кривую иногда называют также *синусоидальной*.

Параметры C, λ, c в уравнении (114) носят следующие названия:

C — амплитуда, λ — частота, c — фаза.

Вследствие произведённого сжатия в λ раз период функции, заданной уравнением (114), равен уже не 2π, а $\frac{2\pi}{\lambda}$. Положив

$$\omega = \frac{2\pi}{\lambda},$$

мы получаем

$$\lambda = \frac{2\pi}{\omega},$$

и уравнению (114) можно также придать вид

$$y = C\sin\frac{2\pi}{\omega}(x - c) \quad (C > 0,\ \omega > 0). \tag{115}$$

Частота и период гармонического колебания обратно пропорциональны, причём их произведение равно 2π:

$$\lambda\omega = 2\pi. \tag{116}$$

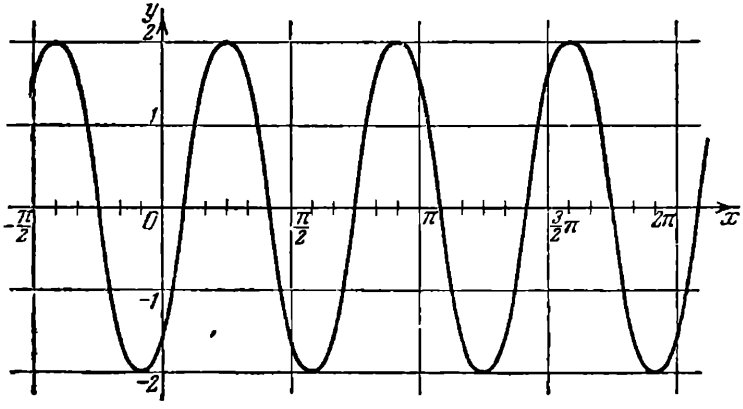


Рис. 51.

На рис. 51 изображено гармоническое колебание

$$y = 2\sin\left(3x - \frac{\pi}{4}\right)$$

с амплитудой $C = 2$, частотой $\lambda = 3$ $\left(\text{периодом } \omega = \frac{2\pi}{3}\right)$ и фазой $c = \frac{\pi}{12}$, причём построение произведено «через четыре точки».

Имея в виду рассмотреть дальше гармонические колебания с одним и тем же периодом, допустим для простоты, что $\omega = 2\pi$, т. е. $\lambda = 1$.

Легко понять, что гармонические колебания вида

$$u = A\cos x \quad \text{и} \quad v = B\sin x \tag{117}$$

(каковы бы ни были знаки A и B) имеют фазы, отличающиеся на четверть периода $\left(\frac{\pi}{2}\right)$.

Всякое гармоническое колебание вида

$$y = C\sin(x - c)$$

может быть разложено на сумму двух таких колебаний:

$$C\sin(x - c) = A\cos x + B\sin x. \tag{118}$$

Чтобы в этом убедиться, достаточно выполнить тригонометрическое преобразование по формуле:

$$C\sin(x - c) = C(\sin x\cos c - \cos x\sin c),$$

и затем положить

$$A = -C\sin c, \quad B = C\cos c. \tag{119}$$

Обратно, *сумма двух гармонических колебаний с одним и тем же периодом представляет собой гармоническое колебание с тем же периодом.*

Доказывая это, сначала допустим, что данные колебания отличаются на четверть периода и имеют вид

$$u = A\cos x, \quad v = B\sin x.$$

Нужно подобрать постоянные $C\,(>0)$ и c таким образом, чтобы удовлетворялось тождество (118). Но тогда дело сводится к решению уравнений (119) относительно неизвестных C и c.

Возводя в квадрат почленно каждое из этих уравнений, складывая и извлекая арифметический корень, мы получаем:

$$C = \sqrt{A^2 + B^2}. \tag{120}$$

Далее из уравнений

$$\begin{cases} \sin c = -\dfrac{A}{\sqrt{A^2 + B^2}}, \\ \cos c = \dfrac{B}{\sqrt{A^2 + B^2}} \end{cases}$$

можно в пределах промежутка $0 \leqslant c < 2\pi$ найти единственное значение c.

Обращаясь к общему случаю, разложим каждое из данных колебаний

$$y_1 = C_1 \sin(x - c_1), \quad y_2 = C_2 \sin(x - c_2)$$

на сумму колебаний вида (118):

$$y_1 = (-C_1 \sin c_1)\cos x + (C_1 \cos c_1)\sin x,$$
$$y_2 = (-C_2 \sin c_2)\cos x + (C_2 \cos c_2)\sin x;$$

тогда получим сумму

$$y = y_1 + y_2 = -(C_1 \sin c_1 + C_2 \sin c_2)\cos x + (C_1 \cos c_1 + C_2 \cos c_2)\sin x,$$

очевидно, вида (118), причём

$$A = -(C_1 \sin c_1 + C_2 \sin c_2), \quad B = C_1 \cos c_1 + C_2 \cos c_2.$$

Сумма простых гармонических колебаний с различными периодами уже не является простым гармоническим колебанием.

Если частоты слагаемых колебаний соизмеримы

$$\lambda_1 = m_1\lambda, \quad \lambda_2 = m_2\lambda,$$

(где m_1 и m_2 — целые положительные числа), то сумма колебаний

$$y_1 = C_1 \sin \lambda_1 (x - c_1) \quad \text{и} \quad y_2 = C_2 \sin \lambda_2 (x - c_2),$$

равная

$$y = C_1 \sin \lambda_1 (x - c_1) + C_2 \sin \lambda_2 (x - c_2),$$

представляет собой так называемое *сложное гармоническое колебание*[1]) с периодом

$$\omega = \frac{2\pi}{\lambda}.$$

(В случае же несоизмеримости частот сумма не является периодической функцией.)

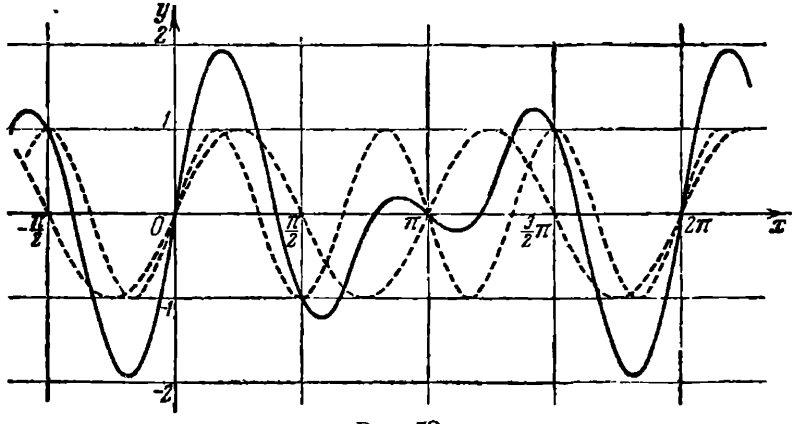


Рис. 52.

На рис. 52 сплошной линией изображена сумма (по точкам) (см. § 3) колебаний с частотами 2 и 3 (изображённых пунктирными линиями):

$$y_1 = \sin 2x \quad \text{и} \quad y_2 = \sin 3x.$$

[1]) Определение см. ниже, § 27.

§ 27. Тригонометрические многочлены

Подобно тому как всякая функция вида

$$f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n$$

(где n — целое положительное число; $c_0, c_1, \ldots, c_n$ — постоянные коэффициенты) носит название *рационального многочлена* (относительно переменной x), введено также особое наименование для функций вида

$$f(x) = a_0 + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \ldots \\ \ldots + (a_n \cos nx + b_n \sin nx), \quad (121)$$

где n — целое положительное число; $a_0, a_1, \ldots, a_n, b_1, b_2, \ldots, b_n$ — постоянные коэффициенты. Такие функции называются *тригонометрическими многочленами* (относительно переменной x). Число n (при условии $a_n^2 + b_n^2 \neq 0$) называется *порядком* тригонометрического многочлена. Коэффициент a_0 есть *свободный член*, выражение $a_1 \cos x + b_1 \sin x$ — *первый* член, выражение $a_2 \cos 2x + b_2 \sin 2x$ — *второй* член тригонометрического многочлена и т. д. График m-го члена многочлена (при $m \geqslant 1$) представляет собой *простое гармоническое колебание частоты* m (т. е. периода $\frac{2\pi}{m}$). График всякого тригонометрического многочлена порядка $n\,(\geqslant 2)$ носит название *сложного гармонического колебания*. Пример такого колебания был указан в предыдущем параграфе.

Так как всякая функция периода ω имеет также периодами все числа, кратные ω, то каждый член многочлена (121) имеет периодом число 2π. Так как, с другой стороны, сумма функций некоторого периода также есть функция с этим самым периодом, то сам многочлен (121) имеет период 2π [1]).

Совершенно очевидно, что

1) сумма двух (или большего числа) тригонометрических многочленов порядка $\leqslant n$ есть также тригонометрический многочлен порядка $\leqslant n$;

2) при умножении тригонометрического многочлена на постоянное число он остаётся тригонометрическим многочленом, без повышения порядка.

---

[1]) Термины «тригонометрический многочлен» и «сложное гармоническое колебание» относятся также к выражениям более общего вида

$$a_0 + \left(a_1 \cos \frac{2\pi x}{\omega} + b_1 \sin \frac{2\pi x}{\omega}\right) + \ldots + \left(a_n \cos \frac{2\pi n x}{\omega} + b_n \sin \frac{2\pi n x}{\omega}\right),$$

возникающим при «растяжении» в $\frac{\omega}{2\pi}$ раз по направлению оси Ox. Такие многочлены, конечно, имеют период ω.

В дальнейшем, однако, ради простоты рассматриваются лишь многочлены вида (121) с периодом 2π.

Докажем следующую теорему.

Всякий рациональный многочлен $P(u, v)$ относительно двух переменных

$$u = \cos x, \quad v = \sin x$$

есть тригонометрический многочлен относительно переменной x. Порядок этого многочлена не превышает степени многочлена $P(u, v)$ относительно пары переменных u, v.

Доказательство разобьём на несколько ступеней.

1) Каждое из выражений вида

$$\cos px \cos qx, \quad \sin px \cos qx, \quad \cos px \sin qx, \quad \sin px \sin qx \qquad (122)$$

(где p и q — целые положительные числа) есть тригонометрический многочлен порядка $p+q$. Это следует из элементарных тождеств:

$$\cos px \cos qx = \frac{1}{2}[\cos(p+q)x + \cos(p-q)x],$$

$$\sin px \cos qx = \frac{1}{2}[\sin(p+q)x + \sin(p-q)x],$$

$$\cos px \sin qx = \frac{1}{2}[\sin(p+q)x - \sin(p-q)x],$$

$$\sin px \sin qx = \frac{1}{2}[-\cos(p+q)x + \cos(p-q)x].$$

2) Произведение двух тригонометрических многочленов порядков r и s есть тригонометрический многочлен порядка $r+s$.

Предположим, что перемножаются многочлены

$$f(x) = a_0 + (a_1 \cos x + b_1 \sin x) + \ldots + (a_r \cos rx + b_r \sin rx)$$

и

$$g(x) = a'_0 + (a'_1 \cos x + b'_1 \sin x) + \ldots + (a'_s \cos sx + b'_s \sin sx).$$

Их произведение есть сумма конечного числа членов вида (122) с постоянными коэффициентами, и следовательно, в силу предварительных замечаний (1) и (2), также есть тригонометрический многочлен. Порядок его, очевидно, не превышает $r+s$, но не может и быть меньше, так как члены порядка $r+s$ получаются только при перемножении выражений $a_r \cos rx + b_r \sin rx$ и $a'_s \cos sx + b'_s \sin sx$, именно, они таковы:

$$\frac{1}{2}\left\{(a_r a'_s - b_r b'_s)\cos(r+s)x + (a_r b'_s + b_r a'_s)\sin(r+s)x\right\}.$$

Коэффициенты при $\cos(r+s)x$ и $\sin(r+s)x$ не могут обратиться в нуль одновременно, так как если $a_r^2 + b_r^2 > 0$, то из уравнений

$$a_r a'_s - b_r b'_s = 0,$$
$$b_r a'_s + a_r b'_s = 0$$

сейчас же следует:

$$a'_s = b'_s = 0.$$

3) В частности, выражения вида

$$\cos^h x \sin^k x \tag{123}$$

(где каждое из чисел h и k есть целое положительное число или нуль) являются тригонометрическими многочленами порядка $h+k$.

В самом деле, каждый из $h+k$ множителей в выражении (123) есть или $\cos x$ или $\sin x$ — тригонометрические многочлены порядка 1.

4) Произвольный рациональный многочлен $P(u, v)$, где $u=\cos x$, $v=\sin x$, есть сумма конечного числа членов вида (123) с постоянными коэффициентами и потому на основании замечаний (1)—(2) есть тригонометрический многочлен. Порядок этого многочлена, очевидно, не превышает наибольшей из сумм $h+k$, т. е. степени многочлена $P(u, v)$. Но этот порядок может быть и меньше (простейший пример: $P(u, v)=u^2+v^2$).

§ 28. Многочлены Чебышева

Мы докажем теперь обратную теорему:

Всякий тригонометрический многочлен порядка n

$$f(x)=a_0+(a_1\cos x+b_1\sin x)+(a_2\cos 2x+b_2\sin 2x)+\ldots$$
$$\ldots+(a_n\cos nx+b_n\sin nx)\qquad (a_n^2+b_n^2\neq 0) \tag{124}$$

представляет собой рациональный многочлен $P(u, v)$ степени n (точно) *относительно пары переменных*

$$u=\cos x,\ v=\sin x.$$

Достаточно установить это по отношению к простейшим тригонометрическим многочленам вида $\cos nx$ и $\sin nx$ (где n — произвольное целое положительное число); в самом деле, при сложении рациональных многочленов и умножении их на постоянные числа, очевидно, снова получаются рациональные многочлены. Останется ещё проверить утверждение, касающееся степени.

Докажем сначала, что существуют такие рациональные многочлены 1) $T_n(u)$ степени n относительно $u=\cos x$ и 2) $U_n(u)$ степени $n-1$ относительно u, которые удовлетворяют тождествам относительно x:

$$\cos nx=T_n(\cos x), \tag{125}$$

$$\sin nx=U_n(\cos x)\sin x. \tag{126}$$

Чтобы уяснить существо вопроса, посмотрим, что́ получается при значениях $n=1, 2, 3$.

Если $n=1$, то нужно положить

$$T_1(u)=u,\ U_1(u)=1. \tag{127}$$

Если $n=2$, то

$$\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1,$$
$$\sin 2x = 2\sin x \cos x = 2\cos x \sin x,$$

так что

$$T_2(u) = 2u^2 - 1, \quad U_2(u) = 2u. \tag{128}$$

Если $n=3$, то аналогично получаем:

$$\begin{aligned}\cos 3x &= \cos(2x+x) = \cos 2x \cos x - \sin 2x \sin x = \\ &= (2\cos^2 x - 1)\cos x - 2\cos x \sin^2 x = \\ &= (2\cos^2 x - 1)\cos x - 2\cos x(1-\cos^2 x) = \\ &= 4\cos^3 x - 3\cos x,\end{aligned}$$

$$\begin{aligned}\sin 3x &= \sin(2x+x) = \sin 2x \cos x + \cos 2x \sin x = \\ &= 2\cos^2 x \sin x + (2\cos^2 x - 1)\sin x = \\ &= (4\cos^2 x - 1)\sin x,\end{aligned}$$

так что

$$T_3(u) = 4u^3 - 3u, \quad U_3(u) = 4u^2 - 1. \tag{129}$$

Таким образом можно продолжать и дальше.

Дальнейшее рассуждение основывается на методе полной индукции. Допустим, что существование многочленов $T_n(u)$ и $U_n(u)$ установлено, и сами они уже вычислены; посмотрим, как установить существование многочленов $T_{n+1}(u)$ и $U_{n+1}(u)$ и как их вычислить.

Мы имеем, пользуясь тождествами (125) и (126), которые предположены доказанными:

$$\begin{aligned}\cos(n+1)x &= \cos(nx+x) = \cos nx \cos x - \sin nx \sin x = \\ &= T_n(\cos x)\cos x - U_n(\cos x)\sin^2 x = \\ &= T_n(\cos x)\cos x - U_n(\cos x)(1-\cos^2 x),\end{aligned}$$

и, с другой стороны,

$$\begin{aligned}\sin(n+1)x &= \sin(nx+x) = \sin nx \cos x + \cos nx \sin x = \\ &= U_n(\cos x)\sin x \cos x + T_n(\cos x)\sin x = \\ &= \{U_n(\cos x)\cos x + T_n(\cos x)\}\sin x.\end{aligned}$$

Так как $T_n(u)$ и $U_n(u)$, по предположению, — многочлены соответственно степеней n и $n-1$, то, очевидно, выражения

$$T_n(u)\,u - U_n(u)(1-u^2) \quad \text{и} \quad U_n(u)\,u + T_n(u)$$

— также многочлены степеней соответственно $n+1$ и n. Обозначая их через $T_{n+1}(u)$ и $U_{n+1}(u)$:

$$\left.\begin{aligned}T_{n+1}(u) &= u\,T_n(u) + (u^2-1)\,U_n(u), \\ U_{n+1}(u) &= T_n(u) + u\,U_n(u),\end{aligned}\right\} \tag{130}$$

мы получаем те тождества, которые нам нужно было доказать:

$$\cos(n+1)x = T_{n+1}(\cos x),$$
$$\sin(n+1)x = U_{n+1}(\cos x)\sin x.$$

Существование многочленов $T_1(u)$ и $U_1(u)$, определяемых формулами (127) и удовлетворяющих тождествам (125) и (126) при $n=1$ было проверено непосредственно. Многочлены $T_n(u)$ и $U_n(u)$ при $n \geqslant 2$ вычисляются по рекуррентным зависимостям (130). Так получаются формулы (127), (128), (129) и следующие:

$$T_4(u) = 8u^4 - 8u^2 + 1, \qquad U_4(u) = 8u^3 - 4u,$$
$$T_5(u) = 16u^5 - 20u^3 + 5u, \qquad U_5(u) = 16u^4 - 12u^2 + 1,$$
$$T_6(u) = 32u^6 - 48u^4 + 18u^2 - 1, \qquad U_6(u) = 32u^5 - 32u^3 + 6u$$

и т. д. Легко проверить с помощью формул (125) и (126), что 1) все многочлены $T_{2n}(u)$ и $U_{2n+1}(u)$ — чётные функции переменной u, а $T_{2n+1}(u)$ и $U_{2n}(u)$ — нечётные, 2) старшие коэффициенты многочленов $T_n(u)$ и $U_n(u)$ равны 2^{n-1}:

$$\left.\begin{aligned} T_n(u) &= 2^{n-1}u^n + \ldots, \\ U_n(u) &= 2^{n-1}u^{n-1} + \ldots \end{aligned}\right\} \qquad (131)$$

Многочлены $T_n(u)$ называются *многочленами Чебышева первого рода*, а $U_n(u)$ — *многочленами Чебышева второго рода*.

Используя многочлены Чебышева, мы можем тригонометрическому многочлену, заданному формулой (124), придать вид

$$f(x) = \{a_0 + a_1T_1(u) + a_2T_2(u) + \ldots + a_nT_n(u)\} + v\{b_1U_1(u) + b_2U_2(u) + \ldots + b_nU_n(u)\}.$$

Выражения, стоящие в первой и во второй фигурных скобках, — рациональные многочлены от переменной u:

$$\left.\begin{aligned} L(u) &= a_0 + a_1T_1(u) + a_2T_2(u) + \ldots + a_nT_n(u), \\ M(u) &= b_1U_1(u) + b_2U_2(u) + \ldots + b_nU_n(u), \end{aligned}\right\} \qquad (132)$$

так что

$$f(x) = L(u) + vM(u). \qquad (133)$$

Если $a_n \neq 0$, то многочлен $L(u)$ — степени n; если $b_n \neq 0$, то многочлен $M(u)$ — степени $n-1$. Но так как, по предположению, хотя бы один из коэффициентов a_n и b_n отличен от нуля, то непременно или $L(u)$ — степени n или $M(u)$ — степени $n-1$. В обоих случаях многочлен $L(u) + vM(u)$ — степени n относительно пары переменных u, v.

Многочлен $L(u)+vM(u)$ удовлетворяет всем требованиям, которые были предъявлены теоремой к многочлену $P(u, v)$. Но он — не единственный, удовлетворяющий этим требованиям [1]).

Мы укажем ещё одну возможную форму многочлена $P(u, v)$.

Заменяя в тождествах (125) и (126) x через $\frac{\pi}{2}-x$, мы получим новые тождества:

$$\left.\begin{aligned}\cos n\left(\frac{\pi}{2}-x\right)&=T_n(\sin x),\\ \sin n\left(\frac{\pi}{2}-x\right)&=U_n(\sin x)\cos x,\end{aligned}\right\}\tag{134}$$

которым, расчленяя случаи n чётного и n нечётного, можно также придать следующий вид:

$$\cos 2nx=(-1)^n T_{2n}(\sin x),\quad \cos(2n+1)x=(-1)^n U_{2n+1}(\sin x)\cos x,$$

$$\sin 2nx=(-1)^{n+1} U_{2n}(\sin x)\cos x,\quad \sin(2n+1)x=(-1)^n T_{2n+1}(\sin x).$$

Отсюда получается:

$$f(x)=\{a_0+b_1T_1(v)-a_2T_2(v)-b_3T_3(v)+a_4T_4(v)+\dots\}+$$
$$+u\{a_1U_1(v)+b_2U_2(v)-a_3U_3(v)-b_4U_4(v)+a_5U_5(v)+\dots\}$$

и, полагая

$$\left.\begin{aligned}L_1(v)&=a_0+b_1T_1(v)-a_2T_2(v)-b_3T_3(v)+a_4T_4(v)+\dots,\\ M_1(v)&=a_1U_1(v)+b_2U_2(v)-a_3U_3(v)-b_4U_4(v)+\dots,\end{aligned}\right\}\tag{135}$$

мы находим новую форму для многочлена $P(u, v)$:

$$f(x)=L_1(v)+uM_1(v).\tag{136}$$

Полезно обратить внимание на несколько следующих частных случаев. Условимся буквой P обозначать рациональный многочлен. Тогда

1) Если $b_1=b_2=b_3=b_4=\dots=0$, т. е. если $f(x)$ есть тригонометрический многочлен вида

$$f(x)=a_0+a_1\cos x+a_2\cos 2x+\dots+a_n\cos nx+\dots,$$

то он представляет собой рациональный многочлен от $\cos x$:

$$f(x)=P(\cos x).$$

2) Если $a_0=a_1=a_2=a_3=\dots=0$, т. е. если $f(x)$ есть тригонометрический многочлен вида

$$f(x)=b_1\sin x+b_2\sin 2x+\dots+b_n\sin nx+\dots,$$

---

[1]) Это следует из того, что если $u=\cos x$, $v=\sin x$, то выражение u^2, встретившееся где-нибудь в формуле, можно заменить через $1-v^2$.

то он может быть представлен в виде

$$f(x) = \sin x \cdot P(\cos x).$$

3) Если $a_1 = a_3 = a_5 = \ldots = 0$, $b_2 = b_4 = b_6 = \ldots = 0$, т. е. если тригонометрический многочлен $f(x)$ имеет вид

$$f(x) = a_0 + b_1 \sin x + a_2 \cos 2x + b_3 \sin 3x + \ldots,$$

то он представляет собой рациональный многочлен от $\sin x$:

$$f(x) = P(\sin x).$$

4) Если $a_0 = a_2 = a_4 = \ldots = 0$ и $b_1 = b_3 = b_5 = \ldots = 0$, т. е. тригонометрический многочлен $f(x)$ имеет вид

$$f(x) = a_1 \cos x + b_2 \sin 2x + a_3 \cos 3x + b_4 \sin 4x + \ldots,$$

то он может быть представлен в виде

$$f(x) = \cos x \cdot P(\sin x).$$

Нетрудно также проверить, что все эти утверждения обратимы.

§ 29. Тангенс и другие дробные тригонометрические функции

Перейдём к рассмотрению дробных тригонометрических функций, понимая под таковыми рациональные функции от основных тригонометрических $\sin x$ и $\cos x$, содержащие по крайней мере одну операцию деления.

Простейшими элементарными дробными тригонометрическими функциями являются такие, которые содержат только одно действие и именно — деление. Их — четыре: тангенс, котангенс, секанс и косеканс; они определяются формулами

$$\operatorname{tg} x = \frac{\sin x}{\cos x}, \quad \operatorname{ctg} x = \frac{\cos x}{\sin x}, \quad \sec x = \frac{1}{\cos x} \quad \text{и} \quad \operatorname{cosec} x = \frac{1}{\sin x}. \tag{137}$$

Тангенс заслуживает особого внимания. С помощью соотношений (110) мы получаем:

$$\operatorname{tg}(x+\pi) = \frac{\sin(x+\pi)}{\cos(x+\pi)} = \frac{-\sin x}{-\cos x} = \frac{\sin x}{\cos x} = \operatorname{tg} x, \tag{138}$$

откуда ясно, что тангенс имеет период π, вдвое меньший, чем синус и косинус. С другой стороны, тангенс — нечётная функция

$$\operatorname{tg}(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{-\sin x}{\cos x} = -\operatorname{tg} x. \tag{139}$$

Таким образом, достаточно изучить поведение тангенса в пределах первой четверти $\left(0 < x < \frac{\pi}{2}\right)$. Имея график тангенса в этих пределах, можно его продолжить, пользуясь симметрией относи-

тельно центра O, для $-\frac{\pi}{2} < x < 0$, а дальше — продолжать вправо и влево по свойству периодичности.

При $x = 0$ получаем:

$$\operatorname{tg} 0 = \frac{\sin 0}{\cos 0} = 0;$$

так как в пределах первой четверти числитель $\sin x$ возрастает (от 0 до 1) и знаменатель $\cos x$ убывает (от 1 до 0), то дробь, определяющая тангенс, возрастает, и именно, от нуля до бесконечности. При $x = \frac{\pi}{2}$ тангенс «не существует», «теряет смысл», «терпит разрыв», так как знаменатель $\cos x$ обращается в нуль, тогда как числитель $\sin x$ равен единице. Таким образом, график тангенса не имеет ни одной точки на прямой $x = \frac{\pi}{2}$.

Тангенс положителен в первой четверти и (по свойству нечётности) отрицателен в четвёртой; по свойству периодичности снова положителен в третьей и отрицателен во второй. Это иллюстрируется схемой:

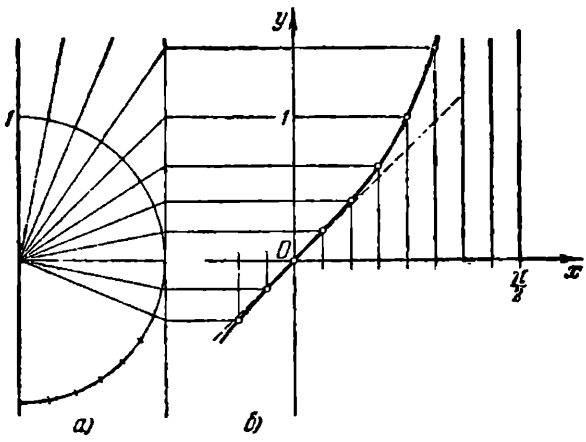
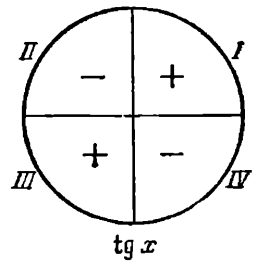


Рис. 53.



На рис. 53 показано точечное построение графика тангенса в пределах первой четверти — посредством деления первой четверти на восемь равных частей. Возвращаясь к единичному кругу (рис. 54), мы видим, что если точка M поставлена в конце дуги AM длины x, то точка P, взятая на пересечении радиус-вектора OM с касательной к кругу в начальной точке A, как раз имеет ординату, равную $\operatorname{tg} x$; в самом деле, из подобия треугольников OAP и OM_1M следует

$$\frac{AP}{OA} = \frac{M_1M}{OM_1},$$

т. е.

$$\frac{AP}{1} = \frac{\sin x}{\cos x}, \quad AP = \operatorname{tg} x.$$

Имея в виду это геометрическое истолкование тангенса, при точечном построении достаточно проектировать точки пересечения радиус-векторов с вертикальной касательной в начальной точке единичного круга (см. рис. 53, *а*) на соответствующие вертикали в плоскости Oxy (см. рис. 53, *б*).

Существенно отметить одну особенность графика тангенса. На рис. 54 дуга AM меньше отрезка AP[1]), и потому отношение $\frac{AP}{AM}$ больше единицы; но когда M приближается к A, это отношение приближается к единице. На рис. 53, *б* этому соответствует следующее: абсцисса каждой точки графика меньше ординаты, и потому отношение ординаты к абсциссе больше единицы; но с приближением точки к началу координат O это отношение приближается к единице. Кривая около начала координат O, будучи расположена выше биссектрисы $y=x$,

$$x<\operatorname{tg}x \quad \left(0<x<\frac{\pi}{2}\right),$$

тесно к ней примыкает (касается её в начале координат O).

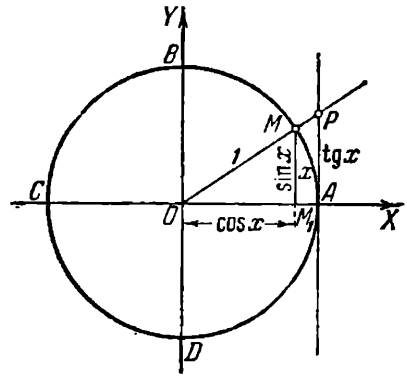


Рис. 54.

Не останавливаясь на более подробном рассмотрении свойств функций котангенс, секанс и косеканс, отметим лишь, что их графики получаются из графиков соответственно тангенса, косинуса и синуса посредством построения графика величины, обратной по отношению к данной функции (см. § 4), в согласии с определениями (137). Но график котангенса из графика тангенса может быть получен ещё другим, более простым, способом: так как

$$\operatorname{tg}\left(\frac{\pi}{2}-x\right)=\frac{\sin\left(\frac{\pi}{2}-x\right)}{\cos\left(\frac{\pi}{2}-x\right)}=\frac{\cos x}{\sin x}=\operatorname{ctg}x,$$

или

$$\operatorname{ctg}x=-\operatorname{tg}\left(x-\frac{\pi}{2}\right),$$

то достаточно над графиком тангенса произвести следующие преобразования:

1) отразить его симметрично относительно оси Ox,

---

[1]) Обосновать это логически можно следующим образом. Обозначим через A' точку на единичном круге, расположенную симметрично точке A относительно прямой OM. Дуга AMA' меньше ломаной APA' (из двух выпуклых линий объемлемая меньше объемлющей); значит, AM меньше AP.

2) перенести параллельно оси Ox на отрезок $\frac{\pi}{2}$.

Нетрудно понять, что совокупность этих преобразований равносильна симметричному отражению относительно оси $x = \frac{\pi}{4}$ [1]).

Подобным же образом график косинуса получается из графика синуса и график косеканса из графика секанса [2]).

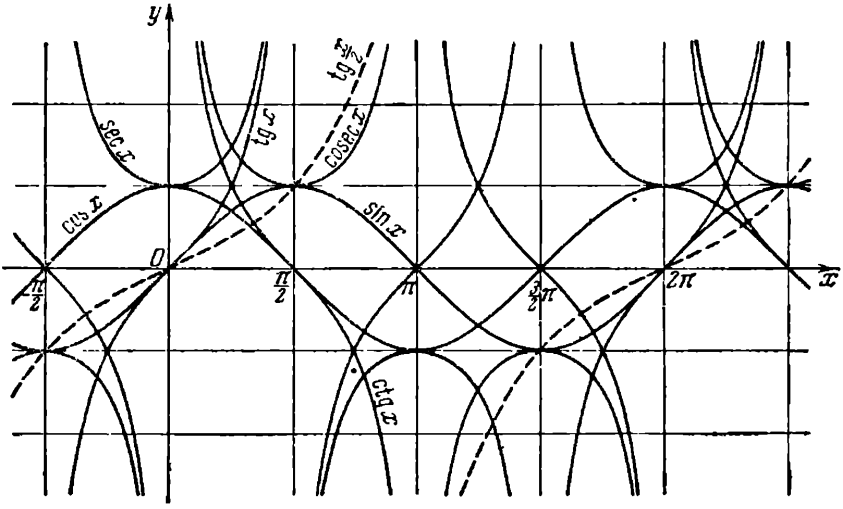


Рис. 55.

На рис. 55 изображены совместно графики шести тригонометрических функций

$$\sin x, \quad \cos x, \quad \operatorname{tg} x, \quad \operatorname{cosec} x, \quad \sec x \quad \text{и} \quad \operatorname{ctg} x.$$

Желая отдать себе отчёт во взаимных связях между шестью тригонометрическими функциями, положим ради краткости

$$\left.\begin{aligned} u &= \cos x, \quad u_1 = \sec x, \\ v &= \sin x, \quad v_1 = \operatorname{cosec} x, \\ w &= \operatorname{tg} x, \quad w_1 = \operatorname{ctg} x. \end{aligned}\right\} \tag{140}$$

Основным соотношениям (101) и (137) можно тогда придать вид

$$\left.\begin{aligned} &1)\ u^2 + v^2 = 1, \quad 3)\ uu_1 = 1, \\ &2)\ w = \frac{u}{v}, \quad\quad 4)\ vv_1 = 1, \\ &\qquad\qquad\ 5)\ ww_1 = 1. \end{aligned}\right\} \tag{141}$$

---

[1]) При этом учитывается центральная симметрия графика тангенса относительно центра O.

[2]) Отсюда термин «кофункции».

Совокупность этих пяти равенств обладает тем свойством, что если задано числовое значение какой-нибудь одной из шести переменных, то, не прибегая ни к каким действиям, кроме четырёх арифметических и извлечения квадратного корня, можно определить значения остальных пяти переменных (решить систему уравнений (141) относительно назначенных пяти букв). Но по смыслу вопроса извлекать квадратный корень приходится алгебраически, т. е. в тех случаях, когда он встречается, остаётся неопределённость в знаке, двузначность. Таким образом, если система соотношений (141) и позволяет рассматривать из шести величин (u, v, w, u_1, v_1, w_1) каждую как функцию каждой, то, как показывает более детальное рассмотрение, из $5 \times 6 = 30$ возникающих при этом функций лишь 6 оказываются однозначными (рациональными), остальные же содержат квадратные корни и потому двузначны. Например, если задано значение синуса $v = \frac{1}{2}$, то для косинуса получается два значения $u = \pm\frac{\sqrt{3}}{2}$, и выбор знака может быть сделан не иначе, как исходя из каких-либо дополнительных данных, позволяющих судить о том, в какой четверти заключено значение независимой переменной x.

Из шести тригонометрических функций нет ни одной, значение которой однозначно определяло бы значение всех остальных. Всё же функция тангенс с этой точки зрения представляет особенный интерес. Если $\operatorname{tg} x = w$, то значение синуса и косинуса вычисляются по формулам

$$\left.\begin{aligned} \sin x &= \frac{w}{\pm\sqrt{1+w^2}}, \\ \cos x &= \frac{1}{\pm\sqrt{1+w^2}}, \end{aligned}\right\} \tag{142}$$

причём знаки перед двумя радикалами могут быть какие угодно, но одинаковые. Замечательно то, что при таких условиях квадраты $\sin^2 x$, $\cos^2 x$, произведение $\sin x \cos x$, а значит, и такие функции, как $\sin 2x$ и $\cos 2x$ через w выражаются уже рационально:

$$\left.\begin{aligned} \sin 2x &= \frac{2w}{1+w^2}, \\ \cos 2x &= \frac{1-w^2}{1+w^2}. \end{aligned}\right\} \tag{143}$$

Теперь легко понять, что, вводя новую функцию с периодом 2π, а именно, тангенс половинного угла

$$t = \operatorname{tg}\frac{x}{2},$$

мы получим, заменяя в формулах (143) x через $\frac{x}{2}$, w через t, или в результате непосредственного вычисления:

$$\left.\begin{aligned}\sin x &= \frac{2t}{1+t^2},\\ \cos x &= \frac{1-t^2}{1+t^2}.\end{aligned}\right\} \tag{144}$$

Таким образом, *синус, косинус, а следовательно, и остальные четыре тригонометрические функции выражаются через тангенс половинного угла рационально.* По значению $\operatorname{tg}\frac{x}{2}$ можно вычислить однозначно значения всех тригонометрических функций.

График функции $t=\operatorname{tg}\frac{x}{2}$ (он изображён на рис. 55 пунктиром) получается из графика $w=\operatorname{tg}x$ посредством растяжения в 2 раза по направлению оси Ox. Из рис. 55 ясно, что каждому значению t соответствует (в пределах периода длиной 2π) одно и только одно значение x, а следовательно, — одно значение каждой из функций u, v, w, u_1, v_1, w_1.

Конечно, функция $t_1=\frac{1}{t}=\operatorname{ctg}\frac{x}{2}$ способна играть такую же роль.

§ 30. Представление функций, рационально зависящих от тригонометрических, через одну или две из них

Всякая рациональная функция от элементарных тригонометрических функций одной и той же переменной x может быть представлена как рациональная функция от двух элементарных функций:

$$u=\cos x \quad \text{и} \quad v=\sin x.$$

Это следует из того обстоятельства, что по формулам (141) функция $w=\operatorname{tg}x$, а также функции $u_1=\sec x$, $v_1=\operatorname{cosec}x$ и $w_1=\operatorname{ctg}x$ выражаются рационально через u и v. При этом нужно принять во внимание, что рациональная функция от одной или нескольких переменных, из которых каждая зависит рационально от одной или большего числа других переменных, сама есть, очевидно, рациональная же функция от этих других переменных.

Пример.

$$\frac{1}{2}\left(\frac{\sec x}{1+\operatorname{tg}x}-\frac{\operatorname{cosec}x}{1-\operatorname{ctg}x}\right)=\frac{v}{u^2-v^2}=\frac{\cos x}{\cos^2 x-\sin^2 x}.$$

I. *Всякая рациональная функция от элементарных тригонометрических функций одной и той же переменной x может быть представлена как рациональная функция от одной лишь функции — тангенса половинного угла:*

$$t=\operatorname{tg}\frac{x}{2}.$$

Это следует из формул (144) предыдущего параграфа.

Пример.

$$\frac{\cos x}{\cos^2 x - \sin^2 x} = \frac{1-t^4}{1-6t^2+t^4} = \frac{1-\operatorname{tg}^4\frac{x}{2}}{1-6\operatorname{tg}^2\frac{x}{2}+\operatorname{tg}^4\frac{x}{2}}.$$

Теорема I носит общий характер. Следующие теоремы II, III и IV содержат в условии дополнительные предположения и выделяют таким образом важные частные случаи.

II. *Если рациональная функция от элементарных тригонометрических функций является* а) *чётной,* б) *нечётной относительно независимой переменной x, то она может быть представлена* а) *как рациональная функция от переменной $u=\cos x$,* б) *как произведение рациональной функции от переменной $u=\cos x$ на $v=\sin x$.*

а) Обозначим рассматриваемую функцию через $f(x)$. По теореме I имеем тождество

$$f(x) = R_1\left(\operatorname{tg}\frac{x}{2}\right)^{1}. \tag{145}$$

С другой стороны, по условию,

$$f(-x) = f(x). \tag{146}$$

Отсюда следует тождество:

$$R_1\left(\operatorname{tg}\frac{(-x)}{2}\right) = R_1\left(\operatorname{tg}\frac{x}{2}\right),$$

т. е.

$$R_1\left(-\operatorname{tg}\frac{x}{2}\right) = R_1\left(\operatorname{tg}\frac{x}{2}\right), \tag{147}$$

или

$$R_1(-t) = R_1(t)^{2}, \tag{148}$$

так что рациональная функция R_1 — чётная.

---

[1]) Буквами R со значками обозначаются дальше различные рациональные функции.

[2]) Тождество (148) следует из тождества (147), так как при любом t можно подобрать x так, чтобы удовлетворялось равенство

$$\operatorname{tg}\frac{x}{2} = t.$$

В таком случае она может быть представлена как рациональная функция от t^2 [1])

$$R_1(t) = R_2(t^2).$$

Но

$$t^2 = \operatorname{tg}^2 \frac{x}{2} = \frac{1-\cos x}{1+\cos x} = \frac{1-u}{1+u}.$$

Поэтому

$$f(x) = R_1(t) = R_2(t^2) = R_2\left(\frac{1-u}{1+u}\right) = R_3(u).$$

б) Если вместо тождества (146) согласно условию теоремы имеется тождество

$$f(-x) \equiv -f(x),$$

то вместо (148) мы получаем:

$$R_1(-t) = -R_1(t),$$

так что функция $R_1(t)$ — нечётная. Но тогда функция $\frac{R_1(t)}{t}$ чётная, и потому, как раньше,

$$\frac{R_1(t)}{t} = R_2(t^2) = R_3(u).$$

---

[1]) *Если $R(x)$ есть чётная рациональная функция от x, то она есть рациональная функция от x^2.*

Эта теорема может показаться очевидной. Вот её доказательство.

Пусть $R(x) = \frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены.

Разделяя в них члены чётной и нечётной степени, напишем

$$P(x) = P_1(x^2) + xP_2(x^2), \quad Q(x) = Q_1(x^2) + xQ_2(x^2),$$

$$R(x) = \frac{P_1(x^2) + xP_2(x^2)}{Q_1(x^2) + xQ_2(x^2)},$$

где P_1, P_2, Q_1 и Q_2 — новые многочлены. Из тождества $R(-x) = R(x)$ следует тождество

$$\frac{P_1(x^2) - xP_2(x^2)}{Q_1(x^2) - xQ_2(x^2)} = \frac{P_1(x^2) + xP_2(x^2)}{Q_1(x^2) + xQ_2(x^2)},$$

или

$$\frac{P_1(x^2)}{Q_1(x^2)} = \frac{P_2(x^2)}{Q_2(x^2)}.$$

В этой пропорции и правая и левая часть есть одна и та же рациональная функция от x^2. Обозначая её через $S(x^2)$, мы получаем:

$$P_1(x^2) = Q_1(x^2)\,S(x^2) \quad \text{и} \quad P_2(x^2) = Q_2(x^2)\,S(x^2),$$

и отсюда следует тождество

$$R(x) = S(x^2).$$

В таком случае

$$f(x) = R_1(t) = tR_3(u) = \operatorname{tg}\frac{x}{2} \cdot R_3(u).$$

Но

$$\operatorname{tg}\frac{x}{2} = \frac{\sin x}{1+\cos x}, \tag{149}$$

и поэтому

$$f(x) = \frac{\sin x}{1+\cos x} R_3(u) = \frac{v}{1+u} R_3(u) = vR_4(u).$$

III. *Если рациональная функция от элементарных тригонометрических функций* а) *не меняется при замене независимой переменной* x *через* $\pi - x$, б) *меняет лишь знак при этой замене, то она может быть представлена* а) *как рациональная функция переменной* $v = \sin x$, б) *как произведение рациональной функции* $v = \sin x$ *на* $u = \cos x$.

а) Тождеству

$$f(x) = f(\pi - x), \tag{150}$$

заменяя x через $x + \frac{\pi}{2}$, можно придать вид

$$f\left(\frac{\pi}{2} + x\right) = f\left(\frac{\pi}{2} - x\right),$$

или

$$f_1(-x) = f_1(x),$$

где положено

$$f_1(x) = f\left(x + \frac{\pi}{2}\right). \tag{151}$$

Итак, функция $f_1(x)$ — чётная. По теореме IIа в таком случае она может быть представлена как рациональная функция от $u = \cos x$:

$$f_1(x) = R_1(u).$$

Другими словами, имеет место тождество

$$f\left(x + \frac{\pi}{2}\right) = R_1(\cos x).$$

Заменяя в нём x через $x - \frac{\pi}{2}$, получаем:

$$f(x) = R_1\left(\cos\left(x - \frac{\pi}{2}\right)\right) = R_1(\sin x) = R_1(v).$$

б) Если вместо тождества (150) имеет место по условию тождество

$$f(\pi - x) = -f(x),$$

то тогда функция $f_1(x)$, определяемая тождеством (151), — нечётная, и потому на основании теоремы IIб можно написать

$$f_1(x) = vR_1(u) = \sin x \cdot R_1(\cos x),$$

или

$$f\left(x + \frac{\pi}{2}\right) = \sin x \cdot R_1(\cos x).$$

Заменяя x через $x - \frac{\pi}{2}$, получаем:

$$f(x) = -\cos x \cdot R_1(\sin x) = \cos x \cdot R_2(\sin x) = uR_2(v).$$

IV. *Если рациональная функция от элементарных тригонометрических функций* а) *не изменяется при замене* x *на* $x + \pi$ (т. е. имеет период π), б) *меняет лишь знак при этой замене, то она может быть представлена* а) *как рациональная функция от* $w = \operatorname{tg} t$, б) *как произведение рациональной функции* w *на* $u = \cos x$ *или* $v = \sin x$.

а) Раз функция $f(x)$ обладает периодом π

$$f(x + \pi) = f(x),$$

то функция

$$f_1(x) = f\left(\frac{x}{2}\right)$$

обладает периодом 2π; в самом деле,

$$f_1(x + 2\pi) = f\left(\frac{x + 2\pi}{2}\right) = f\left(\frac{x}{2} + \pi\right) = f\left(\frac{x}{2}\right) = f_1(x).$$

В таком случае, по теореме I, существует тождество

$$f_1(x) = R_1\left(\operatorname{tg}\frac{x}{2}\right);$$

заменяя в нём x через $2x$, получаем:

$$f(x) = f_1(2x) = R_1(\operatorname{tg} x).$$

б) Если функция $f(x)$ удовлетворяет тождеству

$$f(x + \pi) = -f(x),$$

то каждая из функций

$$f_1(x) = \frac{f(x)}{\cos x} \quad \text{и} \quad f_2(x) = \frac{f(x)}{\sin x}$$

имеет период π. Значит, по теореме IVa имеем:

$$f_1(x) = R_1(w), \quad f_2(x) = R_2(w),$$

откуда

$$f(x) = f_1(x)\cos x = uR_1(w),$$
$$f(x) = f_2(x)\sin x = vR_2(w).$$

Примечание. Легко проверить, что теоремы I, II, III и IV обратимы. Для теорем II, III и IV это устанавливается непосредственно; что касается теоремы I, то достаточно указать на тождество (149).

Таким образом, теоремы I—IV дают условия, необходимые и достаточные для того, чтобы функция могла быть представлена в той или иной из рассмотренных форм.

§ 31. Примеры исследования функций, рационально зависящих от тригонометрических. Тригонометрические уравнения

Теоремы I, II, III и IV, изложенные в предыдущем параграфе, и родственные им теоремы в § 28 (см. 1 и 3) открывают возможности для элементарного исследования функций, рационально зависящих от основных тригонометрических. Если с помощью введения новой независимой переменной — одной из основных тригонометрических функций или тангенса половинной дуги — удаётся свести данную функцию к рациональной функции от новой переменной, то тем самым, поскольку поведение основных тригонометрических функций можно считать известным, имеются основания судить и о поведении данной сложной функции (см. § 5, п. 9).

Те же теоремы полезны и при решении тригонометрических уравнений. Главная трудность при решении уравнений заключается в том, чтобы, «алгебраизируя» их, удачно выбрать новую переменную. Теоремы позволяют сделать выбор по простым формальным признакам, чем обусловливается направление дальнейших тождественных преобразований.

Следует отметить, что если переход к новой переменной $t=\operatorname{tg}\frac{x}{2}$ позволяет *всегда* произвести рационализацию (это, так сказать, «универсальная» подстановка), тем не менее в случае, если возможна одна из подстановок $u=\cos x$, $v=\sin x$ или $w=\operatorname{tg} x$, то вновь получаемая рациональная функция, как правило, оказывается проще; поэтому можно рекомендовать к «универсальной» подстановке прибегать лишь «в крайнем случае».

Обратимся к примерам на исследование функций.

Пример 1.

$$y=f(x)=\frac{1}{3+2\cos x}.$$

Эта фувкция — рациональная относительно $u=\cos x$. При неограиичеииом изменении x переменная u меняется в пределах от -1 до $+1$, и функция $\frac{1}{3+2u}$ в этом промежутке $\left(\text{как и всюду, кроме точки разрыва } u=-\frac{3}{2}\right)$ — убывающая (см. § 5, п. п. 1, 2, 7). Поэтому функция $y=f(x)$ убывает в тех промежутках, где $\cos x$ возрастает, и возрастает в тех промежутках, где он убывает (§ 5, п. 9). Так как функция $f(x)$ чётная, то достаточно рассмотреть

половину периода $0 \leqslant x \leqslant \pi$: на другую половину $-\pi \leqslant x \leqslant 0$ график продолжается по симметрии. В пределах первых двух четвертей $\cos x$ убывает от 1 до -1; значит, функция $f(x)$ возрастает и именно, от $\frac{1}{5}$ до 1.

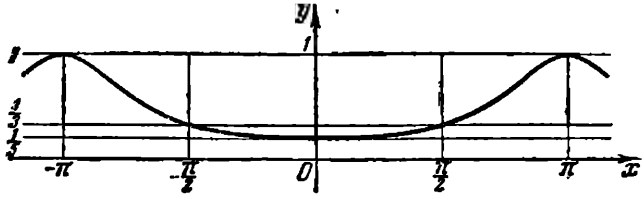


Рис. 56.

При $x=\frac{\pi}{2}$ (положение «равновесия» для $u=\cos x$) значение $f(x)$ равно $\frac{1}{3}$ (рис. 56).

Пример 2.

$$y=f(x)=\frac{\sin^2 x-\sin 2x}{\cos^2 x}.$$

Так как эта функция, очевидно, не меняется при замене x на $x+\pi$ (имеет период π), то естественно ввести переменную $w=\operatorname{tg} x$. Преобразование даёт:

$$y=\operatorname{tg} x(\operatorname{tg} x-2)=w(w-2).$$

Функция

$$w(w-2)=(w-1)^2-1$$

убывает в промежутке $-\infty<w<1$ и возрастает в промежутке $1<w<\infty$.

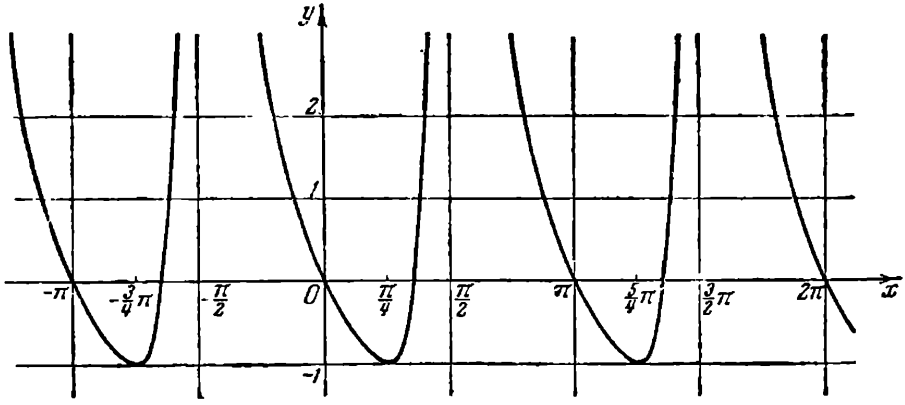


Рис. 57.

Функция $\operatorname{tg} x$ —возрастающая; значение $w=1$ она принимает (в промежутке периода $0<x<\pi$) при $x=\frac{\pi}{4}$; при $x=\frac{\pi}{2}$ она терпит разрыв. Когда x возрастает от 0 до $\frac{\pi}{4}$, то w возрастает от нуля до 1, а y убывает от 0 до -1;

когда x возрастает от $\frac{\pi}{4}$ до $\frac{\pi}{2}$, то w возрастает от 1 до бесконечности, а y — от -1 до бесконечности (значение $y=0$ получается при $w=2$, т. е. при $x \approx 63\frac{1}{2}°$); когда x возрастает, далее, от $\frac{\pi}{2}$ до π, то w (после разрыва) возрастает от $-\infty$ до 0, и тогда y убывает от $+\infty$ до 0. Заметим ещё, что при $x=\frac{3}{4}\pi$ мы имеем $w=-1$, $y=3$ (рис. 57).

Пример 3.

$$y=f(x)=\frac{1}{2}(3\cos x-\cos 3x).$$

Так как функция $f(x)$ — чётная, то достаточно исследовать полупериод $0 \leqslant x \leqslant \pi$, взяв в качестве вспомогательной переменной $u=\cos x$. Так как

$$\cos 3x=4\cos^3 x-3\cos x=4u^3-3u$$

(многочлен Чебышева $T_3(u)$), то

$$y=\frac{1}{2}[3u-(4u^3-3u)]=3u-2u^3.$$

Получившийся многочлен третьей степени сведём к «стандартной форме» (см. § 10, (β)) заменой

$$u=z\sqrt{\frac{3}{2}},\quad z=u\sqrt{\frac{2}{3}},$$

и тогда будем иметь

$$y=-3\sqrt{\frac{3}{2}}(z^3-z).$$

Поведение функции z^3-z меняется в точках $z=\pm\sqrt{\frac{1}{3}}$ (см. § 10), которым соответствуют точки $u=\pm\sqrt{\frac{1}{2}}$ и, далее, $x=\frac{\pi}{4}$ и $x=\frac{3}{4}\pi$.

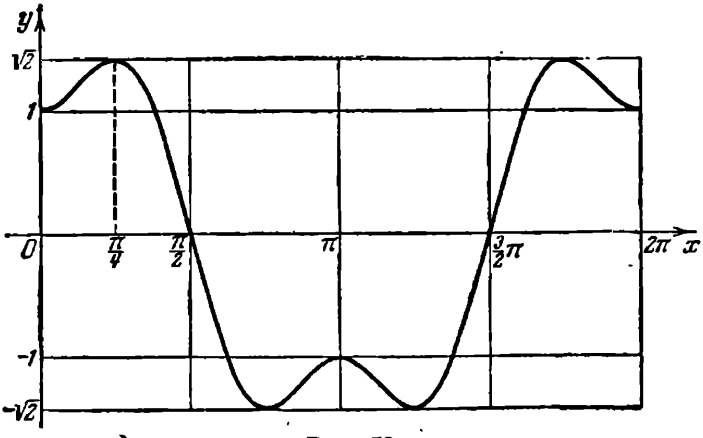


Рис. 58.

Рассматривая образованные этими точками промежутки, мы убеждаемся, пользуясь общими положениями об изменении сложных функций (см. § 5), в справедливости следующей таблицы (см. рис. 58):

x	u	z	z^3-z	y
возрастает от 0 до $\frac{\pi}{4}$	убывает от 1 до $\sqrt{\frac{1}{2}}$	убывает от $\sqrt{\frac{2}{3}}$ до $\sqrt{\frac{1}{3}}$	убывает от $-\frac{1}{3}\sqrt{\frac{2}{3}}$ до $-\frac{2}{3}\sqrt{\frac{1}{3}}$	возрастает от 1 до $\sqrt{2}$
возрастает от $\frac{\pi}{4}$ до $\frac{3}{4}\pi$	убывает от $\sqrt{\frac{1}{2}}$ до $-\sqrt{\frac{1}{2}}$	убывает от $\sqrt{\frac{1}{3}}$ до $-\sqrt{\frac{1}{3}}$	возрастает от $-\frac{2}{3}\sqrt{\frac{1}{3}}$ до $\frac{2}{3}\sqrt{\frac{1}{3}}$	убывает от $\sqrt{2}$ до $-\sqrt{2}$
возрастает от $\frac{3}{4}\pi$ до π	убывает от $-\sqrt{\frac{1}{2}}$ до -1	убывает от $-\sqrt{\frac{1}{3}}$ до $-\sqrt{\frac{2}{3}}$	убывает от $\frac{2}{3}\sqrt{\frac{1}{3}}$ до $\frac{1}{3}\sqrt{\frac{2}{3}}$	возрастает от $-\sqrt{2}$ до -1

Сразу видно, что при $x=\frac{\pi}{2}$ мы получаем $y=0$.

Полезно в этом примере сопоставить результаты проведённого исследования с точечным построением, заключающимся в составлении «полусуммы» графиков

$$y_1=\cos 3x \quad \text{и} \quad y_2=-3\cos x.$$

Пример 4.

$$y=f(x)=\frac{\cos 2x}{\sin x}.$$

Эта функция — нечётная, так что достаточно рассмотреть полупериод $0\leqslant x\leqslant\pi$; график можно затем продолжить, пользуясь симметрией относительно центра O. Функция $f(x)$ не изменяется при замене x на $\pi-x$; отсюда видно, что естественно взять в качестве вспомогательной переменной $v=\sin x$. Вместе с тем отмеченное обстоятельство показывает, что прямая $x=\frac{\pi}{2}$ является осью симметрии графика; таким образом, достаточно рассмотреть даже четверть периода $0\leqslant x\leqslant\frac{\pi}{2}$: на вторую четверть график продолжается по симметрии.

Мы получаем:

$$y=\frac{\cos^2 x-\sin^2 x}{\sin x}=\frac{1-2\sin^2 x}{\sin x}=\frac{1}{\sin x}-2\sin x=\frac{1}{v}-2v, \qquad (152)$$

или же

$$y=2\frac{\left(\sqrt{\frac{1}{2}}+v\right)\left(\sqrt{\frac{1}{2}}-v\right)}{v}. \qquad (153)$$

С увеличением v выражение $\frac{1}{v}-2v$ убывает (всюду, кроме точки разрыва $v=0$), так как $\frac{1}{v}$ убывает, а $2v$ возрастает.

Когда x возрастает от 0 до $\frac{\pi}{2}$, то $v=\sin x$ возрастает от 0 до 1, и тогда $y=\frac{1}{v}-2v$ убывает от $+\infty$ до -1. Из формулы (153) ясно, что

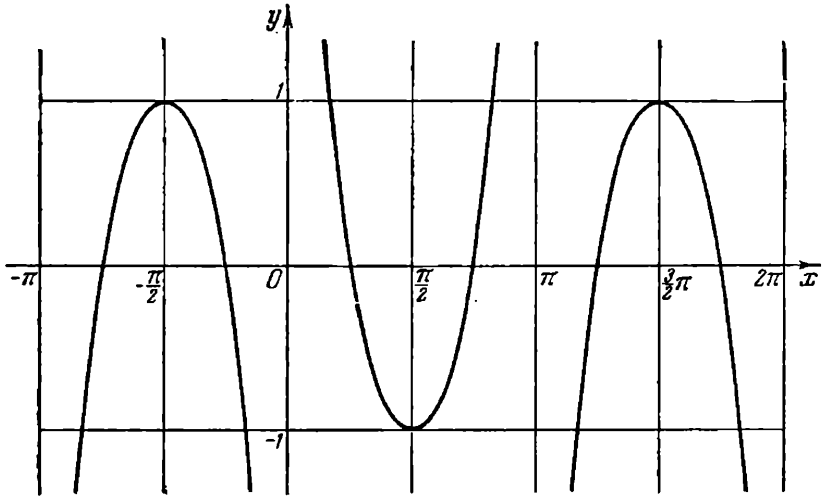


Рис. 59.

значение нуль y принимает (в первой четверти) при $v=\sqrt{\frac{1}{2}}$, т. е. при $x=\frac{\pi}{4}$ (рис. 59).

Пример 5.

$$y=f(x)=\frac{1+\cos x}{3-\sin x}.$$

В этом примере не удаётся обнаружить элементов симметрии или периодичности (помимо периода 2π). Исследуем функцию посредством замены

$$t=\operatorname{tg}\frac{x}{2}.$$

Выражая $\cos x$ и $\sin x$ через t по формулам (144), мы получаем:

$$y=\frac{1+\frac{1-t^2}{1+t^2}}{3-\frac{2t}{1+t^2}}=\frac{2}{3-2t+3t^2} \tag{154}$$

и, далее (см. § 9),

$$y=\frac{2}{3}\cdot\frac{1}{\left(t-\frac{1}{3}\right)^2+\frac{8}{9}}. \tag{155}$$

Поведение выражения $\left(t-\frac{1}{3}\right)^2+\frac{8}{9}$ (в зависимости от t) меняется при $t=\frac{1}{3}$, чему соответствует значение x, равное корню ξ уравнения

$$\operatorname{tg}\frac{x}{2}=\frac{1}{3},$$

а именно,

$$x=\xi\sim 37^\circ.$$

Когда x возрастает от $-\pi$ до ξ, то t возрастает от $-\infty$ до $\frac{1}{3}$, выражение $\left(t-\frac{1}{3}\right)^2+\frac{8}{9}$ убывает от ∞ до $\frac{8}{9}$ и, значит, y возрастает от 0 до $\frac{3}{4}$. Когда же x возрастает от ξ до π, то t возрастает от $\frac{1}{3}$ до ∞, выражение $\left(t-\frac{1}{3}\right)^2+\frac{8}{9}$ возрастает от $\frac{8}{9}$ до ∞ и y убывает от $\frac{3}{4}$ до 0.

Заметим, кроме того, что при $x=0$, $x=\frac{\pi}{2}$ и $x=-\frac{\pi}{2}$ получается соответственно: $y=\frac{2}{3}$, $y=\frac{1}{2}$ и $y=\frac{1}{4}$ (рис. 60).

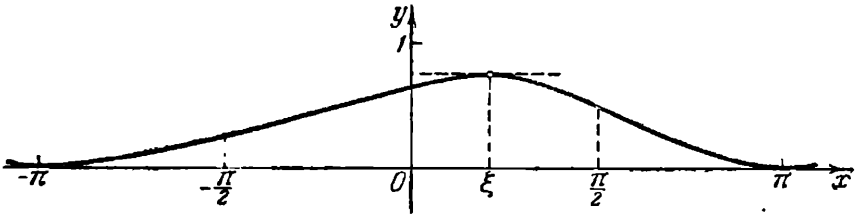


Рис. 60.

Переходя к примерам на решение тригонометрических уравнений, уместно заметить, что при решении уравнения нас непосредственно интересует только один вопрос из общего плана исследования (см. § 5), именно: *при каких значениях переменной x данная функция принимает значение нуль?* Так как при решении этого вопроса существенно лишь разложение на множители данного выражения, то одинаково удобно прибегать к теоремам II — IVa и к теоремам II — IVб предыдущего параграфа (или к 1—4, § 28). Характер использования этих теорем таков, что, установив, какую тригонометрическую функцию удобно взять в качестве вспомогательной переменной, нетрудно дать надлежащее направление выполняемым тождественным преобразованиям.

Пример 6.

$$9-11\cos x+13\cos 2x-3\cos 3x=0.$$

Так как левая часть — чётная функция x, то (на основании теоремы IIа) можно ввести переменную $u=\cos x$. Принимая во внимание, что

$$\cos 2x=2u^2-1,\quad \cos 3x=4u^3-3u$$

(многочлены Чебышева 1-го рода, см. стр. 108), находим:

$$9-11\cos x+13\cos 2x-3\cos 3x=-2(6u^3-13u^2+u+2).$$

Выделив множитель $u-2$ многочлена третьей степени, мы легко разложим затем на множители получившийся трёхчлен второй степени

$$6u^3-13u^2+u+2=(u-2)(6u^2-u-1)=(u-2)(2u-1)(3u+1).$$

Итак, уравнению можно придать вид

$$(\cos x-2)(2\cos x-1)(3\cos x+1)=0.$$

Множитель $\cos x-2$ никогда не обращается в нуль; значит, уравнение удовлетворяется лишь при условии, что

$$\cos x=\frac{1}{2} \quad \text{или} \quad \cos x=-\frac{1}{3},$$

откуда получаются два корня

$$x_1=\frac{\pi}{3}, \quad x_2=-\frac{\pi}{3},$$

и нетрудно найти два остальных (в пределах периода).

Пример 7.

$$3\sin 2x+4\cos 3x-3\sin 4x=0.$$

При замене x на $\pi-x$ левая часть уравнения меняет знак. Поэтому на основании теоремы III б, вводим переменную $v=\sin x$. Тогда с использованием многочленов Чебышева обоих родов получаем:

$\sin 2x=\cos x\cdot 2v$,

$\cos 3x=4\cos^3 x-3\cos x=\cos x(4\cos^2 x-3)=\cos x\cdot(1-4v^2)$,

$\sin 4x=(8\cos^3 x-4\cos x)\sin x=\cos x(8\cos^2 x-4)\sin x=(4v-8v^3)\cos x$;

после подстановки уравнение принимает вид

$$(12v^3-8v^2-3v+2)\cos x=0.$$

Многочлен третьей степени удаётся разложить на множители:

$$(4v^2-1)(3v-2)\cos x=0.$$

или

$$(4\sin^2 x-1)(3\sin x-2)\cos x=0,$$

и мы легко находим восемь корней уравнения (в пределах периода):

$$x_1=\frac{\pi}{2}, \quad x_2=\frac{3}{2}\pi, \quad x_3=\frac{\pi}{4}, \quad x_4=\frac{3}{4}\pi, \quad x_5=\frac{5}{4}\pi, \quad x_6=\frac{7}{4}\pi;$$

два последних корня определяются из уравнения

$$\sin x=\frac{2}{3}.$$

Пример 8. Требуется найти корни уравнения

$$1+\sin x\cos x=35\cos^4 x$$

в пределах первой четверти.

Так как обе части уравнения содержат лишь члены чётных степеней относительно $\cos x$ и $\sin x$ и, следовательно, левая и правая его части не меняются при замене x на $x+\pi$, то имеет смысл ввести переменную $w=\operatorname{tg} x$. Пользуясь формулами (142), приводим уравнение к виду

$$(1+w^2)(1+w+w^2)=35. \tag{156}$$

В первой четверти $w = \operatorname{tg} x > 0$, и потому нам нужны лишь положительные корни уравнения (156). Его левая часть — возрастающая функция w и, значит, уравнение (156) имеет не более одного положительного корня. Легко проверить, что один корень есть: это $w = 2$.

Остаётся найти значение x в первой четверти из уравнения

$$\operatorname{tg} x = 2.$$

§ 32. Обратные [1]) тригонометрические функции

В уравнении

$$y = \sin x \tag{157}$$

поменяем местами x и y; получим

$$x = \sin y. \tag{158}$$

Записывают также равенство (158) в виде

$$y = \operatorname{Arcsin} x; \tag{159}$$

равенства (159) и (158) следует рассматривать, таким образом, как равносильные.

Обратная тригонометрическая функция $\operatorname{Arcsin} x$ (*арксинус* x буквально — «дуга, синус которой равен x») не является однозначной. Вернёмся к единичному кругу, с помощью которого была определена функция синус (см. рис. 44).

Каждой данной дуге AM соответствует одно определённое значение синуса—ордината точки M, вертикальный отрезок M_1M. Пусть, обратно, в качестве независимой переменной взят некоторый вертикальный отрезок M_1M (он может быть направлен вверх или вниз, но по длине не должен превосходить единицы). Обозначим его через x:

$$M_1M = x.$$

Тогда на вопрос, какая дуга имеет этот отрезок своим синусом, — однозначного ответа дать нельзя. Таких дуг существует бесконечное множество: требуемым свойством обладает не только дуга AM, но и дуга ABN, а также $ABCDAM$ и $ABCDABN$ и т. д. и ещё «отрицательные» дуги $ADCN$, $ADCBM$ и т. д. Итак, если функция $\operatorname{Arcsin} x$ имеет значение y, то она имеет значение $\pi - y$, а также (вообще) $y + 2k\pi$ и $(\pi - y) + 2k\pi$.

Свойство неоднозначности функции $\operatorname{Arcsin} x$ ясно видно и из её графика. Мы знаем (см. § 4), что график уравнения (159) получается из графика уравнения (157) посредством симметричного отражения относительно биссектрисы $y = x$ (рис. 61). Из графика функции $\operatorname{Arcsin} x$ видно, что при условии $|x| \leqslant 1$ существует бесконеч-

[1]) См. § 4.

ное множество точек, имеющих данную абсциссу x: если ординату одной из них (всё равно которой) обозначить через y, то ординаты всех охватываются формулами

$$y + 2k\pi,$$
$$(\pi - y) + 2k\pi.$$

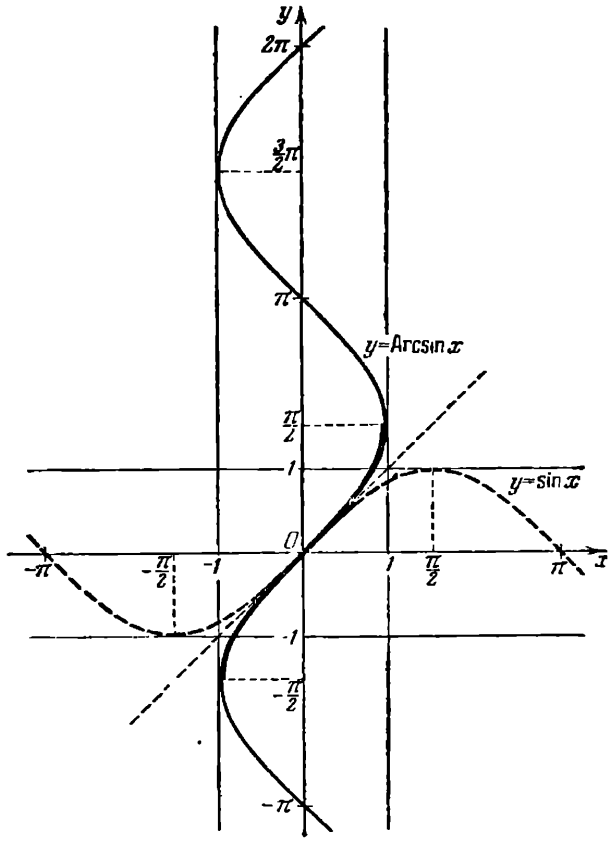


Рис. 61.

Подобно тому как в случае квадратного корня (функция, обратная функции «квадрат») из числа x при условии $x \geqslant 0$ необходимо различать двузначный «алгебраический» корень от однозначного «арифметического», так же точно и в случае арксинуса (функция, обратная функции «синус») от независимой переменной x при условии $|x| \leqslant 1$ — всякий раз, когда могут возникнуть сомнения, — приходится указывать, имеется ли в виду какое-нибудь (любое, безразлично какое) значение рассматриваемой функции или же некоторое определённое, и какое именно. Но различие — в том, что функция $\sqrt{x}$ двузначна, тогда как функция $\operatorname{Arcsin} x$ бесконечно многозначна.

Те значения $y = \operatorname{Arcsin} x$, которые удовлетворяют неравенству $-\frac{\pi}{2} \leqslant y \leqslant +\frac{\pi}{2}$, нередко называют *главными* и обозначают через $\arcsin x$. Таким образом, можно написать

$$\operatorname{Arcsin} x = \begin{cases} \arcsin x + 2k\pi, \\ (\pi - \arcsin x) + 2k\pi. \end{cases} \tag{160}$$

На рис. 61 главные значения арксинуса отмечены жирной линией. Функция $y = \arcsin x$ определена в пределах $-1 \leqslant x \leqslant +1$, притом она — однозначная, нечётная и возрастающая (см. § 52).

В данной статье, выяснив суть дела на одном примере арксинуса, нет необходимости входить в подробности, касающиеся функций, обратных прочим элементарным тригонометрическим функциям. Ограничимся поэтому краткими указаниями и графическими иллюстрациями.

Функция *арккосинус*

$$y = \operatorname{Arccos} x, \tag{161}$$

обратная функции косинус

$$y = \cos x, \tag{162}$$

определена и притом бесконечно многозначна (как и функция $\operatorname{Arcsin} x$) в промежутке $(-1 \leqslant x \leqslant +1)$. Её «главное значение»

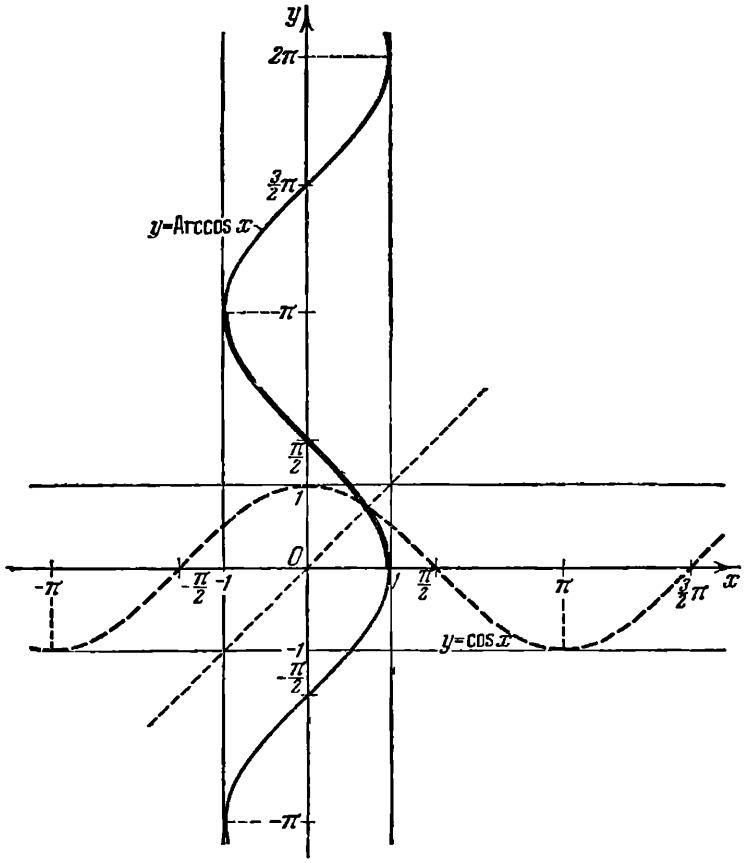


Рис. 62.

$y = \arccos x$ удовлетворяет неравенству $0 \leqslant y \leqslant \pi$; общая формула имеет вид

$$\operatorname{Arccos} x = \pm \arccos x + 2k\pi. \tag{163}$$

Функция $y = \arccos x$ — однозначная и убывающая — определена в промежутке $-1 \leqslant x \leqslant +1$ (рис. 62).

Функция *арктангенс*

$$y = \operatorname{Arctg} x, \tag{164}$$

обратная функции тангенс

$$y = \operatorname{tg} x, \tag{165}$$

определена и притом бесконечно многозначна для всех значений независимой переменной ($-\infty < x < +\infty$). «Главное значение» $\operatorname{arctg} x$ удовлетворяет неравенству $-\frac{\pi}{2} < y < \frac{\pi}{2}$; общая формула имеет вид

$$\operatorname{Arctg} y = \operatorname{arctg} x + k\pi.$$

Функция $y = \operatorname{arctg} x$ — однозначная, нечётная и возрастающая в пределах $-\infty < x < +\infty$. Функция арктангенс никогда не принимает значений вида $\frac{\pi}{2} + k\pi$. График её состоит из бесчисленного множества отдельных «ветвей», переходящих одна в другую при перенесениях параллельно оси Oy на отрезки вида $k\pi$ (рис. 63).

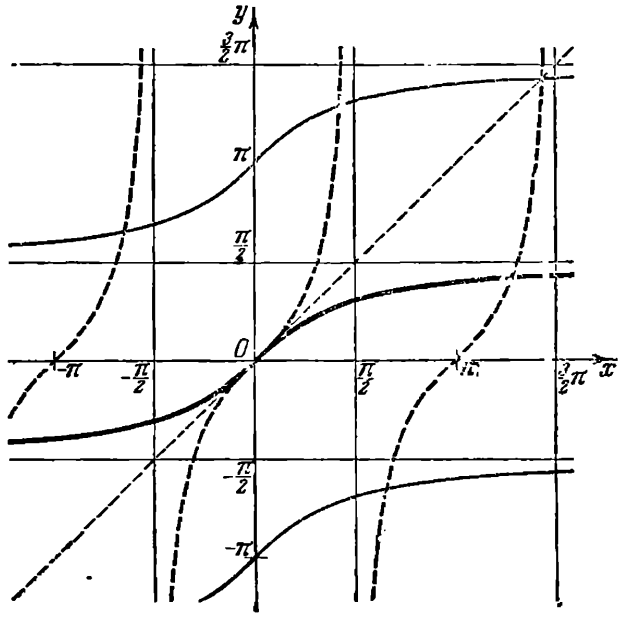


Рис. 63.

Аналогично определяются функции *арккосеканс, арксеканс и арккотангенс*

$$\left.\begin{aligned} y &= \operatorname{Arccosec} x, \\ y &= \operatorname{Arcsec} x, \\ y &= \operatorname{Arcctg} x, \end{aligned}\right\} \tag{166}$$

обратные функциям косеканс, секанс и котангенс

$$\left.\begin{aligned} y &= \operatorname{cosec} x, \\ y &= \sec x, \\ y &= \operatorname{ctg} x. \end{aligned}\right\} \tag{167}$$

Их употребления всегда можно избежать, принимая во внимание, что

$$\left.\begin{aligned} \operatorname{Arccosec} x = \operatorname{Arcsin}\frac{1}{x}\,^{1}), \quad \operatorname{Arcsec} x = \operatorname{Arccos}\frac{1}{x}, \\ \operatorname{Arcctg} x = \operatorname{Arctg}\frac{1}{x}. \end{aligned}\right\} \tag{168}$$

---

[1]) «Дуги, имеющие косеканс, равный x, те же самые, что и дуги, имеющие синус, равный $\frac{1}{x}$» и т. п.

Задачи элементарной математики редко приводят к необходимости пользоваться обратными тригонометрическими функциями в явной форме. Например, задача «вычислить значение $\operatorname{Arcsin} x$ при $x=\frac{1}{3}$» равносильна задаче: «найти все дуги (или углы), синус которых равен $\frac{1}{3}$». В разного рода задачах геометрического содержания удачный выбор неизвестной или переменной, как правило, позволяет обойтись без обратных функций. Напротив, в высшей математике (в интегральном исчислении) обратные круговые функции появляются вполне естественным прямым путём (см. стр. 367) и избегать их было бы крайне неудобно.

Тригонометрические функции связаны между собой большим количеством различных соотношений, значительная часть которых приводится в учебниках тригонометрии; некоторые из формул подобного рода настолько важны в теоретическом и практическом отношениях, что занимающийся математикой запоминает их прочно и навсегда [1]). Обратные функции чрезвычайно обогащают формульный аппарат тригонометрии. Но пользоваться соотношениями, содержащими обратные функции, приходится на самом деле не очень часто и не очень много; и именно математическая практика указывает, какие из подобного рода формул заслуживают преимущественного внимания. С этой точки зрения представляют интерес формулы лишь некоторых типов.

I. *Тригонометрическая функция от обратной тригонометрической* (не обязательно — соответствующего наименования) *есть алгебраическая функция*, именно, выражающаяся через арифметические операции и квадратные радикалы.

Если прямая и обратная функции соответствуют в смысле наименования, то, как явствует из определения, они «погашают» друг друга:

$$\sin(\operatorname{Arcsin} x)=x,\quad \cos(\operatorname{Arccos} x)=x,\quad \operatorname{tg}(\operatorname{Arctg} x)=x\ ^{2}).$$

Рассмотрим теперь случай, когда такого соответствия нет; возьмём, например, выражение $\sin\operatorname{Arccos} x$. Можно написать сразу

$$\sin\operatorname{Arccos} x=\sqrt{1-x^2}, \tag{169}$$

ссылаясь на то, что «синус дуги, косинус которой равен x, есть $\sqrt{1-x^2}$, так как сумма квадратов синуса и косинуса есть 1». Или

---

[1]) Сюда относятся, например, «теоремы сложения» (синус и косинус суммы двух дуг).

[2]) Спросим себя: «Кто отец сына, у которого отец Иван?» Сомнения нет: отец — Иван.

станем рассуждать подробнее: обозначим через y какую-нибудь дугу, косинус которой равен x,

$$\operatorname{Arccos} x = y,$$

тогда, по определению,

$$x = \cos y$$

и, следовательно,

$$\sin \operatorname{Arccos} x = \sin y = \sqrt{1 - \cos^2 y} = \sqrt{1 - x^2}\ ^{1}).$$

Такого же характера соображения приводят в итоге к следующей табличке формул [2]):

$$\sin \operatorname{Arctg} x = \frac{x}{\sqrt{1+x^2}}, \quad \cos \operatorname{Arcsin} x = \sqrt{1-x^2},$$

$$\cos \operatorname{Arctg} x = \frac{1}{\sqrt{1+x^2}}, \quad \operatorname{tg} \operatorname{Arcsin} x = \frac{x}{\sqrt{1-x^2}},$$

$$\sin \operatorname{Arccos} x = \sqrt{1-x^2},$$

$$\operatorname{tg} \operatorname{Arccos} x = \frac{\sqrt{1-x^2}}{x}.$$

II. *Обратная тригонометрическая функция от прямой* (с соответствием наименований или без соответствия) *есть бесконечно многозначное выражение, которое при несоответствии наименований может быть упрощено лишь в случае «кофункций».*

Мы получаем, во-первых [3]):

$$\operatorname{Arcsin}(\sin x) = \begin{cases} x + 2k\pi, \\ (\pi - x) + 2k\pi, \end{cases}$$

$$\operatorname{Arccos}(\cos x) = \pm x + 2k\pi,$$

$$\operatorname{Arctg}(\operatorname{tg} x) = x + k\pi.$$

Мы получаем, во-вторых:

$$\operatorname{Arcsin}(\cos x) = \left(\frac{\pi}{2} \mp x\right) + 2k\pi, \qquad \operatorname{Arccos}(\sin x) = \pm\left(\frac{\pi}{2} - x\right) + 2k\pi,$$

и т. д.

---

[1]) Конечно, радикалы здесь и ниже следует понимать в алгебраическом смысле: выбор знака может быть сделан лишь в том случае, если известно, в какой четверти находится дуга, определяемая обратной круговой функцией.

[2]) Если дуга, определяемая обратной функцией в левой части формул, расположенных одна над другой, одна и та же, то радикалы в правой части, конечно, должны иметь одинаковые знаки.

[3]) Спросим: «Какой сын у отца, у которого сын Иван?» Не известно: может быть, Иван, а может быть — один из его братьев.

В самом деле, например $\operatorname{Arcsin}(\cos x)$ обозначает всякую дугу, синус которой равен $\cos x$. Так как

$$\cos x = \sin\left(\frac{\pi}{2} - x\right),$$

то $\frac{\pi}{2} - x$ есть одна из таких дуг; всякая иная из этих дуг или в сумме с нею составляет π, или отличается от той или другой на величину, кратную 2π.

III. Нередко приходится *одну из обратных тригонометрических функций заменять другой,* с надлежащим изменением аргумента. Пусть, например, одно из значений $\operatorname{Arcsin} x$ требуется представить как арккосинус от некоторой величины; обозначим последнюю буквой y:

$$\operatorname{Arcsin} x = \operatorname{Arccos} y.$$

Из написанного равенства следует:

$$y = \cos(\operatorname{Arcsin} x),$$

или по формуле п. I

$$y = \sqrt{1 - x^2}.$$

Итак,

$$\operatorname{Arcsin} x = \operatorname{Arccos} \sqrt{1 - x^2}, \qquad (170)$$

причём эту формулу нужно понимать в том смысле, что каждое значение $\operatorname{Arcsin} x$ (при условии $|x| \leqslant 1$) равно некоторому значению $\operatorname{Arccos} \sqrt{1 - x^2}$; какому именно — подлежит уточнению в зависимости от конкретных обстоятельств.

Подобным же образом

$$\operatorname{Arcsin} x = \operatorname{Arctg} \frac{x}{\sqrt{1 - x^2}},$$

$$\operatorname{Arctg} x = \operatorname{Arcsin} \frac{x}{\sqrt{1 + x^2}} \qquad \text{и т. п.}$$

§ 33. Исследование многочленов Чебышева. Их минимальное свойство

Идея элементарного исследования тригонометрических многочленов (или вообще функций, рационально зависящих от тригонометрических) заключается в том, чтобы посредством тригонометрической подстановки свести вопрос к более простому исследованию рациональной функции.

В немногих случаях возможно осуществить обратный ход мысли: исследование рационального многочлена (или рациональной функции) свести к более простому исследованию тригонометрического многочлена или функции, рационально зависящей от тригонометрических. Дело, конечно, в том, что немногие тригонометрические многочлены успешно поддаются непосредственному исследованию. К числу таких многочленов можно, однако, отнести $\cos nx$ и $\sin nx$.

В качестве примера рассмотрим поведение *многочлена Чебышева первого рода* $T_n(x)$ в промежутке значений независимой переменной от -1 до $+1$. Формуле, его определяющей,

$$\cos nx = T_n(\cos x),$$

заменив $\cos x$ через u, можно придать вид

$$T_n(u) = \cos n \operatorname{Arccos} u. \tag{171}$$

Здесь всё равно, какое значение $\operatorname{Arccos} u$ иметь в виду: функция $\cos nx$ не меняется при замене x на $-x$ или на $x + 2k\pi$. Для простоты допустим, что речь идёт о главных значениях $\operatorname{Arccos} u$.

Предположим сначала, что n — число чётное: $n = 2m$.

Когда u возрастает от -1 до $+1$, то $x = \arccos u$ убывает от π до 0, и значит, величина $nx = n \arccos u$ убывает от $n\pi$ до 0. Но раз эта величина, являющаяся аргументом под знаком косинуса, убывает от $n\pi = 2m\pi$ до 0, то легко понять (представим себе график косинуса), что сам косинус, т. е. многочлен $T_n(u)$, при этом $\frac{n}{2} = m$ раз совершит полное колебание от $+1$ к -1 и обратно. Нетрудно понять, что при этом:

1) Значение $+1$ многочлен $T_n(u)$ принимает в тех точках, где $n \arccos u = 2k\pi$, т. е. $u = \cos \frac{2k\pi}{n}$:

$$T_n\left(\cos \frac{2k\pi}{n}\right) = 1 \qquad \left(0 \leqslant k \leqslant \frac{n}{2}\right). \tag{172}$$

2) Значение -1 он принимает в тех точках, где $n \arccos u = \pi + 2k\pi$, т. е. $u = \cos \frac{(2k+1)\pi}{n}$:

$$T_n\left(\cos \frac{(2k+1)\pi}{n}\right) = -1 \qquad \left(0 \leqslant k < \frac{n}{2}\right). \tag{173}$$

3) Значение 0 он принимает в точках, где $n \arccos u = \frac{\pi}{2} + k\pi$, т. е. $u = \cos \frac{\left(k + \frac{1}{2}\right)\pi}{n}$:

$$T_n\left(\cos \frac{\left(k + \frac{1}{2}\right)\pi}{n}\right) = 0 \qquad (0 \leqslant k \leqslant n - 1). \tag{174}$$

Предполагая, что n — нечётное: $n = 2m + 1$, мы получим такие же результаты, с той разницей, что при увеличении u от -1 до $+1$ выражение $n \arccos u$ убывает от $n\pi = (2m+1)\pi$ до 0 и поэтому $T_n(u)$ сначала возрастает от -1 до $+1$, затем делает ещё m полных колебаний от $+1$ до -1 и обратно. Формулы же (172—174) сохраняются без изменений.

На рис. 64 воспроизведён график многочлена Чебышева

$$y = T_{12}(x)\ ^{1}),$$

при $n = 12$ в пределах, указанных неравенством $-1 \leqslant x \leqslant +1$.

Самым замечательным свойством многочлена Чебышева

$$y = T_n(x)$$

является то, что в промежутке $-1 \leqslant x \leqslant +1$ при убывании x от 1 до -1 он принимает, начиная с правого конца ($x = 1$), естественно, поочерёдно, наибольшие значения $+1$ и наименьшие -1, все равные между собой по абсолютному значению, причём совершается n переходов от $+1$ к -1 или обратно; таким образом, на левом конце ($x = -1$) принимается значение $+1$ или -1 в зависимости от чётности n. За пределами промежутка $|T_n(x)|$ быстро возрастает.

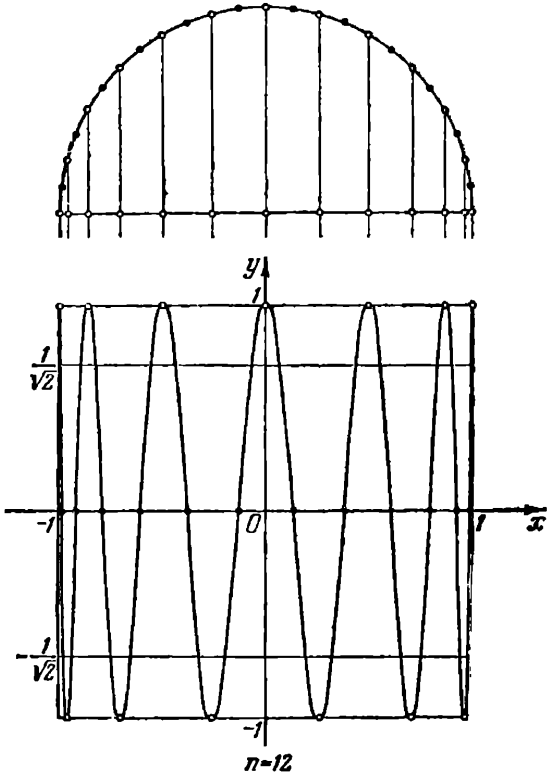


Рис. 64.

Сделаем теперь «сжатие» в 2^{n-1} раз по направлению оси Oy и рассмотрим новый многочлен

$$y = \frac{1}{2^{n-1}} T_n(x). \quad (175)$$

Свойства его графика — такие же, что и графика $T_n(x)$, с той разницей, что абсолютные величины максимумов и минимумов на этот раз равны $\frac{1}{2^{n-1}}$. Вместе с тем, так как старший коэффициент $T_n(x)$ равен 2^{n-1} (см. § 28, (131)), то старший коэффициент нового многочлена равен 1. Итак, среди многочленов степени n, со старшим коэффициентом, равным единице,

$$P_n(x) = x^n + \ldots \quad (176)$$

---

[1]) Для удобства независимое переменное — аргумент многочлена T_n — здесь и дальше обозначено, как обычно, буквой x.

существует такой $\left(\text{именно, } \frac{1}{2^{n-1}} T_n(x)\right)$, который в промежутке $-1 \leqslant x \leqslant 1$ удовлетворяет неравенству

$$|P_n(x)| \leqslant \frac{1}{2^{n-1}}.$$

Мы покажем сейчас, что ещё уменьшить указанную границу нельзя: многочлен $P_n(x)$ степени n вида (176) не может

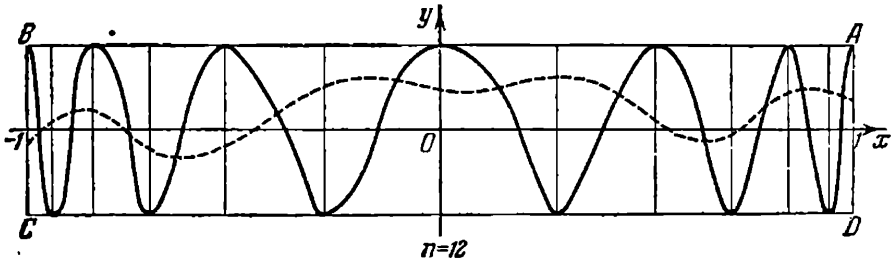


Рис. 65.

при всех значениях x ($|x| \leqslant 1$) удовлетворять неравенству

$$|P_n(x)| < \frac{1}{2^{n-1}}. \tag{177}$$

На рис. 65 сплошной линией показана схематически часть графика (175), заключённая в прямоугольнике с вершинами

$$A\left(+1, \frac{1}{2^{n-1}}\right), \quad B\left(-1, \frac{1}{2^{n-1}}\right), \quad C\left(-1, -\frac{1}{2^{n-1}}\right) \text{ и } D\left(1, -\frac{1}{2^{n-1}}\right).$$

Вертикалями, проведёнными через точки максимума и минимума, этот прямоугольник разбит на n частей — также прямоугольников с одной и той же высотой и различными основаниями. Если бы существовал многочлен $P_n(x)$ вида (176), удовлетворяющий неравенству (177), то часть графика

$$y = P_n(x), \tag{178}$$

ограниченная пределами $-1 \leqslant x \leqslant +1$, была бы заключена внутри полосы, ограниченной прямыми AB и DC (пунктирная кривая на рис. 65); так как она соединяла бы отрезки BC и AD, то этот график имел бы по меньшей мере n общих точек с графиком (175). В самом деле, на пути от BC к AD график (178) должен был бы пересечь все промежуточные вертикали, и внутри каждого частного прямоугольника было бы по меньшей мере по одной общей точке графиков (178) и (175), так как, очевидно, *не могут не пересечься лежащие внутри прямоугольника две кривые, из которых одна соединяет противоположные вершины, а другая — внутренние точки противоположных сторон этого прямоугольника.*

Но абсциссы n точек пересечения графиков (178) и (175) (несомненно все — различные) должны быть корнями уравнения

$$\frac{1}{2^{n-1}} T_n(x) - P_n(x) = 0. \tag{179}$$

А это невозможно, так как вследствие равенства старших коэффициентов рассматриваемых многочленов уравнение (179) — степени ниже чем n.

Мы пришли к противоречию; следовательно, сделанное нами допущение о существовании многочлена $P_n(x)$ было неверно.

Доказанное свойство многочленов Чебышева формулируют обыкновенно так: *из всех многочленов $P_n(x)$ степени n вида* (176) *наименее всех «отклоняется от нуля» в промежутке многочлен Чебышева* $\frac{1}{2^{n-1}}T_n(x)$. При этом под «отклонением от нуля» некоторого многочлена $P(x)$ в данном промежутке $a \leqslant x \leqslant b$ следует понимать наибольшее значение $|P(x)|$ в этом промежутке, иначе говоря, — наибольшую (по абсолютной величине) из ординат соответствующего графика.

Это свойство оказалось отправным пунктом созданной П. Л. Чебышевым (в середине прошлого столетия) теории наилучшего приближения функций.

Не останавливаясь подробнее на многочленах Чебышева второго рода (см. § 28, (131))

$$U_n(x) = \frac{\sin n x}{\sin x} = \frac{\sin n \arccos x}{\sqrt{1-x^2}},$$

исследование которых также может быть проведено элементарным методом [1]), рассмотрим ещё подобный же пример дробной рациональной функции, исследование которой сводится к исследованию дробной тригонометрической функции.

Функция, о которой идёт речь, имеет вид

$$y = R_n(x) = \operatorname{tg}(n \operatorname{arctg} x)\ ^{2}). \tag{180}$$

Что эта функция — рациональная, устанавливается методом полной математической индукции. Пусть $R_n(x)$ — рациональная. Тогда

$$R_{n+1}(x) = \operatorname{tg}[(n+1)\operatorname{arctg} x] = \operatorname{tg}(n \operatorname{arctg} x + \operatorname{arctg} x) =$$
$$= \frac{\operatorname{tg}(n \operatorname{arctg} x) + \operatorname{tg}(\operatorname{arctg} x)}{1 - \operatorname{tg}(n \operatorname{arctg} x) \cdot \operatorname{tg}(\operatorname{arctg} x)} = \frac{R_n(x) + x}{1 - R_n(x) \cdot x},$$

откуда видно, что функция $R_{n+1}(x)$ — также рациональная. Но $R_1(x) = \operatorname{tg}(\operatorname{arctg} x) = x$ — функция рациональная; значит, и все функции $R_n(x)$ — рациональные. Например:

$$R_2(x) = \frac{2x}{1-x^2}, \quad R_3(x) = \frac{3x - x^3}{1 - 3x^2}, \quad R_4(x) = \frac{4(x - x^3)}{1 - 6x^2 + x^4} \text{ и т. д.}$$

---

[1]) Максимумы и минимумы многочлена $U_n(x)$ лежат уже не на горизонтальных прямых $y = \pm 1$, а на кривых

$$y = \pm \frac{1}{\sqrt{1-x^2}}.$$

[2]) Выбор значения арктангенса безразличен, так как функция $\operatorname{tg} nx$ не изменяется при замене x на $x + \pi$. Поэтому можно считать, что выбираются «главные значения» арктангенса.

Так как $\operatorname{arctg} x$ — функция возрастающая и $\operatorname{tg} x$ также (кроме точек разрыва), то $R_n(x)$ — тоже функция, возрастающая всюду, кроме точек разрыва. При возрастании x от $-\infty$ до $+\infty$ функция $\operatorname{arctg} x$ возрастает от $-\frac{\pi}{2}$ до $+\frac{\pi}{2}$, а $n\operatorname{arctg} x$ — от $-\frac{n\pi}{2}$ до $+\frac{n\pi}{2}$. Точки разрыва возникают при $n\operatorname{arctg} x=\frac{\pi}{2}+k\pi$, т. е. при $x=\operatorname{tg}\frac{\frac{\pi}{2}+k\pi}{n}$, причём целые числа k должны удовлетворять неравенству

$$-\frac{n\pi}{2}\leqslant\frac{\pi}{2}+k\pi\leqslant+\frac{n\pi}{2},$$

или

$$-\frac{n+1}{2}\leqslant k\leqslant\frac{n-1}{2}. \tag{181}$$

Функция $R_n(x)$ обращается в нуль при условии $n\operatorname{arctg} x=k\pi$, т. е. при $x=\operatorname{tg}\frac{k\pi}{n}$, причём k удовлетворяют неравенству

$$-\frac{n\pi}{2}\leqslant k\pi\leqslant+\frac{n\pi}{2},$$

или

$$-\frac{n}{2}\leqslant k\leqslant\frac{n}{2}. \tag{182}$$

По поводу значений k, в точности равных границам, указываемым соотношениями (181) и (182), следует заметить, что 1) «разрыв при $x=\pm\infty$»

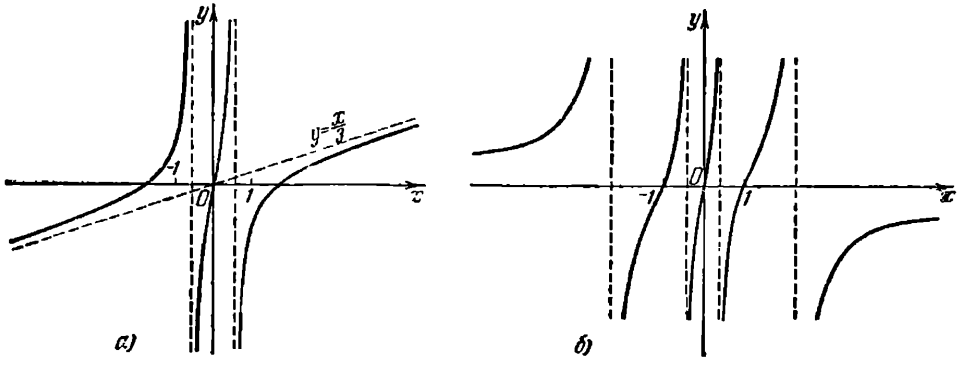


Рис. 66.

даёт наклонную асимптоту, 2) «обращение в нуль при $x=\pm\infty$» соответствует случаю, когда ось Ox становится асимптотой.

На рис. 66, *а* и *б* показаны соответственно графики $R_3(x)$ и $R_4(x)$.

ГЛАВА III

ПРЕДЕЛЫ ЧИСЛОВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И ПРЕДЕЛЫ ФУНКЦИЙ

§ 34. Конечные и бесконечные числовые последовательности

Из всякого данного множества (совокупности) E каких угодно объектов можно образовывать последовательности.

Последовательность строится следующим образом. Указывается некоторый объект A_1 из E, называемый *первым* членом последовательности; затем указывается объект A_2, который непосредственно следует за A_1 и называется *вторым* членом; далее, указывается объект A_3, который непосредственно следует за A_2 и называется *третьим* членом, и т. д. Объекты A_1, A_2, A_3 и т. д. — не обязательно различные: среди них могут быть и одинаковые. Процесс построения последовательности заключается в том, что если уже указан некоторый «n-й член», получивший «порядковый номер», или индекс n, то указывается «непосредственно следующий» за ним «$(n+1)$-й член» с индексом $n+1$.

Такой процесс может закончиться на некотором объекте A_N, получившем индекс, равный натуральному числу N: это произойдёт в том случае, если не будет указано никакого объекта, непосредственно следующего за объектом A_N. Тогда N-й член последовательности A_N называется её *последним* членом; индекс N в этом случае обозначает число членов последовательности. Сама последовательность тогда называется *конечной*.

Перечисляя члены конечной последовательности в порядке следования, обычно их разделяют запятыми:

$$A_1,\ A_2,\ A_3,\ \ldots\ ,\ A_N. \tag{1}$$

Многоточие здесь обозначает пропуск членов с индексами, бо́льшими чем 3, но меньшими чем N.

Но процесс построения последовательности можно представлять себе и неограниченно продолжающимся, не имеющим конца. В таком случае не будет существовать последнего элемента последовательности: каково бы ни было натуральное число n, за членом последовательности A_n, имеющим индекс n, будет непосредственно сле-

довать член A_{n+1}, имеющий индекс, на единицу больший. Если мы говорим, что в результате построения получается бесконечная последовательность, то это означает, что, каково бы ни было натуральное число N, всегда найдётся член последовательности, имеющий индекс N; вместе с тем придётся сделать допущение, что всякий член последовательности имеет в качестве индекса некоторое натуральное число [1]).

Члены числовых последовательностей в дальнейшем будут обозначаться маленькими буквами. Самый простой способ задать конечную числовую последовательность заключается в том, чтобы написать все её члены один за другим, в порядке следования:

$$a_1,\ a_2,\ a_3,\ \ldots,\ a_N. \qquad (2)$$

Вот несколько примеров конечных числовых последовательностей:

1) 1, 2, 3, 4, 5 $\quad (N=5)$,
2) 7, 2, 10 $\quad (N=3)$,
3) 1, 2, 3, 1, 2, 3, 1, 2 $\quad (N=8)$.

В случае, если число членов N велико, непосредственное выписывание всех членов становится затруднительным. Иногда, чтобы задать последовательность, прибегают к аналитическому методу.

---

[1]) Нужно объяснить, в чём смысл последней оговорки. Рассмотрим, например, совокупность различных между собой чисел вида

$$1-\frac{1}{n},\quad 1+\frac{1}{n} \text{ и } 2-\frac{1}{n},$$

где n может принимать все натуральные значения, кроме 1 и 2. Если все эти числа расположены в порядке возрастания, то возникает следующая «упорядоченная» система чисел:

$$\frac{2}{3},\ \frac{3}{4},\ \frac{4}{5},\ \frac{5}{6},\ \cdots,\ \frac{7}{6},\ \frac{6}{5},\ \frac{5}{4},\ \frac{4}{3},\ \frac{5}{3},\ \frac{7}{4},\ \frac{9}{5},\ \frac{11}{6},\ \frac{13}{7},\ \cdots$$

В этой системе имеется первый член $\frac{2}{3}$, которому не предшествует непосредственно никакой другой; помимо того, каждому числу непосредственно предшествует и за каждым членом непосредственно следует один и только один член; таким образом, последнего члена в системе не существует. Система, однако, не образует бесконечной последовательности: действительно, приписывая члену $\frac{2}{3}$ индекс 1, члену $\frac{3}{4}$ индекс 2 и т. д., вообще члену вида $1-\frac{1}{n}$ индекс $n-2$, мы «израсходовали» бы все натуральные индексы на члены вида $1-\frac{1}{n}$, и на члены вида $1+\frac{1}{n}$ или $2-\frac{1}{n}$ индексов бы «нехватило».

Располагая все числа данной совокупности в ином порядке (не в порядке возрастания), мы могли бы получить последовательность чисел, например:

$$\frac{2}{3},\ \frac{4}{3},\ \frac{5}{3},\ \frac{3}{4},\ \frac{5}{4},\ \frac{7}{4},\ \frac{4}{5},\ \frac{6}{5},\ \frac{9}{5},\ \frac{5}{6},\ \frac{7}{6},\ \frac{11}{6},\ \cdots$$

Предположим, что существует элементарная функция $f(x)$, не теряющая смысла при целых положительных значениях x в пределах от $x=1$ до $x=N$ и обладающая тем свойством, что при подстановке $x=1$ получается как раз первый член последовательности a_1, при подстановке $x=2$ — второй член и т. д. до $x=N$. Тогда достаточно указать функцию $f(x)$ и число членов последовательности N; воспроизведение всей последовательности

$$f(1),\ f(2),\ f(3),\ \ldots,\ f(N)$$

в этом случае не представит труда.

Например:

1) если $f(x)=x^2$, $N=5$, то получим последовательность

$$1,\ 4,\ 9,\ 16,\ 25;$$

2) если $f(x)=x^3-6x^2+11x-6$, $N=4$, то получим последовательность

$$0,\ 0,\ 0,\ 6;$$

3) если $f(x)=(-1)^{x+1}$, $N=8$, то получим последовательность

$$1,\ -1,\ 1,\ -1,\ 1,\ -1,\ 1,\ -1.$$

В случае какой угодно конечной последовательности, содержащей N членов, существует функция $f(x)$, удовлетворяющая поставленным требованиям [1]).

Бесконечную последовательность объектов записывают следующим образом:

$$A_1,\ A_2,\ A_3,\ \ldots, \tag{3}$$

или

$$A_1,\ A_2,\ A_3,\ \ldots,\ A_n,\ \ldots \tag{4}$$

Первое многоточие в последней строчке означает, как и раньше, пропуск конечного числа членов; но многоточия, стоящие в конце строки, согласно общепринятому условию, всегда обозначают пропуск бесконечного числа членов или, лучше сказать, возможность неограниченного продолжения.

Ради сокращения письма пользуются также весьма часто записью

$$\{A_n\}. \tag{5}$$

---

[1]) Такова, например, для последовательности (2) функция, представляющаяся в виде рационального многочлена степени $N-1$

$$f(x)=\sum_{m=1}^{N}(-1)^{N-m}\frac{(x-1)(x-2)\ldots(x-m+1)(x-m-1)\ldots(x-N)}{(m-1)!(N-m)!}a_m$$

(интерполяционная формула Лагранжа).

Член последовательности A_n, индекс которого обозначен буквой (в данном случае n), называется *общим членом* последовательности.

В настоящей главе будут рассматриваться только числовые последовательности, т. е. последовательности чисел (исключительно — действительных), в следующей главе — последовательности функций (*функциональные* последовательности). Позднее придётся встречаться с последовательностями, составленными и из иных объектов.

Как задать (определить) бесконечную числовую последовательность? Наиболее естественным является аналитический способ, указывающий в явной форме, какие действия нужно выполнить над значком n, чтобы получить общий член последовательности:

$$a_n = f(n) \qquad (n = 1,\ 2,\ 3,\ \ldots).$$

В каждом данном случае подбор функции $f(x)$ к заданной числовой последовательности может представлять бо́льшие или меньшие трудности. Этот вопрос стоит в стороне от целей, преследуемых настоящей статьёй, и ему здесь не будет уделено места. Заметим здесь лишь следующее. Для каждой числовой последовательности функцию $f(x)$ можно видоизменить различными способами. Так, в примере 3 вместо $f(x) = (-1)^{x+1}$ можно взять $f(x) = \cos \pi x$; в любом примере вместо $f(x)$ можно взять $f(x) + \sin \pi x$ и т. п.

Отметим ещё один способ задать последовательность: это способ рекуррентных зависимостей[1]). Он указывает, какие действия нужно произвести над членами последовательности, уже вычисленными

$$a_1,\ a_2,\ \ldots,\ a_n,$$

чтобы получить следующий член a_{n+1}; помимо того, должны быть заданы, в том или ином числе (смотря по характеру зависимости), несколько первых членов («начальные данные»).

Рассмотрим ряд примеров последовательностей, которые будем предполагать (ради простоты записи) бесконечными.

1. Арифметическая прогрессия определяется рекуррентной зависимостью (разностным уравнением)

$$a_{n+1} - a_n = d \qquad (n = 1,\ 2,\ 3,\ \ldots) \tag{6}$$

с начальным данным

$$a_1 = a.$$

Прогрессия имеет вид

$$a,\ a + d,\ a + 2d,\ \ldots,\ a + (n-1)\,d,\ \ldots$$

Формула для общего члена

$$a_n = a + (n-1)\,d \qquad (n = 1,\ 2,\ 3,\ \ldots)$$

задаёт ту же прогрессию аналитически. Правая часть этой формулы зависит от индекса n линейно.

---

[1]) Вместо «рекуррентная зависимость» говорят также «разностное уравнение».

2. Если общий член последовательности a_n задаётся формулой вида

$$a_n = P(n),$$

где через $P(x)$ обозначен некоторый многочлен (целая рациональная функция) степени m относительно x, то последовательность $\{a_n\}$ носит название *арифметической прогрессии порядка m*. Обыкновенной арифметической прогрессии соответствует, таким образом, порядок 1.

Приведём примеры арифметических прогрессий высших порядков:

$$\{n^2\} \equiv 1,\ 4,\ 9,\ 16,\ 25,\ \ldots \qquad \text{(порядка 2)},$$

$$\left\{\frac{n(n-1)}{2}\right\} \equiv 0,\ 1,\ 3,\ 6,\ 10,\ \ldots \qquad \text{(то же)},$$

$$\left\{\frac{n(n^2-1)}{6}\right\} \equiv 0,\ 1,\ 4,\ 10,\ 20,\ \ldots \qquad \text{(порядка 3)}.$$

Легко понять, что если $P(x)$ есть многочлен степени m, то

$$P(x+1) - P(x)$$

есть многочлен степени $m-1$. В самом деле, из формулы

$$P(x) = ax^m + bx^{m-1} + \ldots + l \qquad (a \neq 0)$$

следует, что

$$P(x+1) - P(x) = max^{m-1} + \left[\frac{m(m-1)}{2}a + (m-1)b\right]x^{m-2} + \ldots,$$

и здесь $ma \neq 0$. Отсюда ясно, что «*последовательность разностей*»

$$\{a_{n+1} - a_n\},$$

составленная из арифметической прогрессии порядка m, есть прогрессия порядка $m-1$.

Можно было бы также доказать, что «*последовательность сумм*»

$$\{a_1 + a_2 + \ldots + a_n\},$$

составленная из арифметической прогрессии порядка m, есть прогрессия порядка $m+1$ (но мы не будем приводить здесь доказательства).

Эти свойства легко проверяются на приведённых выше примерах. Они могут также быть использованы при «продолжении» прогрессии вправо. Например, написав под пятью ранее выписанными членами прогрессии второго порядка $\left\{\frac{n(n-1)}{2}\right\}$ соответствующие разности, мы обнаруживаем, что эти разности образуют обыкновенную прогрессию (первого порядка); продолжая её и затем суммируя, получаем продолжение данной прогрессии:

$$0,\ 1,\ 3,\ 6,\ 10,\ 15,\ 21,\ 28,\ \ldots$$

$$1,\ 2,\ 3,\ 4\ \ 5,\ \ 6,\ \ 7,\ \ldots$$

Таким образом, члены арифметической прогрессии любого порядка можно вычислять последовательно с помощью одних сложений.

3. Геометрическая прогрессия определяется рекуррентной зависимостью

$$\frac{a_{n+1}}{a_n} = q \qquad (n = 1,\ 2,\ 3,\ \ldots) \tag{7}$$

с начальным данным

$$a_1 = a.$$

Прогрессия имеет вид

$$a,\ aq,\ aq^2,\ \ldots,\ aq^{n-1},\ \ldots$$

Формула для общего члена

$$a_n = aq^{n-1} \qquad (n = 1,\ 2,\ 3,\ \ldots)$$

задаёт ту же прогрессию аналитически.

4. Последовательность «факториалов» определяется рекуррентной зависимостью

$$a_{n+1} = (n+1)\,a_n \qquad (n = 1,\ 2,\ 3,\ \ldots)$$

с начальным данным

$$a_1 = 1.$$

Последовательность имеет вид

$$1,\ 2,\ 6,\ 24,\ 120,\ \ldots,\ n!,\ \ldots,$$

причём общий член даётся формулой

$$a_n = n! = 1 \cdot 2 \cdot 3 \ldots n.$$

5. Последовательность Фибоначчи[1]) определяется тем условием, что каждый член (начиная с третьего) равен сумме двух предыдущих:

$$a_{n+2} = a_n + a_{n+1} \qquad (n = 3,\ 4,\ \ldots). \tag{8}$$

В качестве начальных данных указываются значения двух первых членов, например,

$$a_1 = 0, \quad a_2 = 1. \tag{9}$$

Тогда без всякого труда по рекуррентной формуле (8) мы находим:

$$a_3 = 0 + 1 = 1, \quad a_4 = 1 + 1 = 2, \quad a_5 = 1 + 2 = 3,$$
$$a_6 = 2 + 3 = 5, \quad a_7 = 3 + 5 = 8, \quad a_8 = 5 + 8 = 13,$$
$$a_9 = 8 + 13 = 21, \quad a_{10} = 13 + 21 = 34$$

и т. д. Написав некоторое число членов последовательности Фибоначчи

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ \ldots, \tag{10}$$

естественно поинтересоваться, нельзя ли задать эту же последовательность аналитически, представляя общий её член a_n в явной форме как функцию значка n (в «замкнутой форме»[2]).

---

[1]) Леонард Пизанский (XIII век).

[2]) Смысл последнего иногда употребляемого оборота речи заключается в том, что в формуле, выражающей a_n через n и содержащей лишь элементарные операции, число этих операций не должно зависеть от n. Например, относительно последовательности факториалов (п. 4) неправильно было бы сказать, что её общий член $a_n = n!$ записан «в замкнутой форме» (хотя и весьма кратко).

Мы приведём здесь такого рода решение «разностного» уравнения (8), предварительно заметив, что существует только одна последовательность, удовлетворяющая требованиям (8) и (9): в самом деле, раз первые два члена заданы, третий определяется однозначно по первым двум, четвёртый — по второму и третьему и т. д. Общий член ряда Фибоначчи, как легко проверить, имеет вид

$$a_n = \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1}\right\} \qquad (n = 1, 2, \ldots).$$

Читателю, которого поразит эта формула, мы сделаем всё же намёк по поводу того, каким путём можно к ней придти. Попытаемся, сначала оставляя в стороне начальные условия (9), найти геометрическую прогрессию, которая обладала бы рекуррентным свойством (8). Полагая в уравнении (8) $a_n = aq^{n-1}$, мы получим

$$aq^{n+1} = aq^{n-1} + aq^n,$$

что при допущении $a \neq 0$, $q \neq 0$ приводит нас дальше к квадратному уравнению

$$q^2 = 1 + q. \tag{11}$$

Решая его, мы убеждаемся, что знаменатель прогрессии должен равняться одному из чисел

$$q = \frac{1+\sqrt{5}}{2} \quad \text{или} \quad q' = \frac{1-\sqrt{5}}{2};$$

обратно, каковы бы ни были числа a и a', прогрессии

$$\left.\begin{array}{l} a,\ aq,\ aq^2,\ aq^3,\ \ldots,\ aq^{N-1},\ \ldots \\ \text{и} \\ a',\ a'q',\ a'q'^2,\ a'q'^3,\ \ldots,\ a'q'^{N-1},\ \ldots \end{array}\right\} \tag{12}$$

удовлетворяют соотношению (8).

Далее, обратим внимание на то обстоятельство, что если каждая из последовательностей (12) удовлетворяет соотношению (8), то «последовательность-сумма»

$$a + a',\ aq + a'q',\ aq^2 + a'q'^2,\ aq^3 + a'q'^3,\ \ldots,\ aq^{N-1} + a'q'^{N-1}$$

также ему удовлетворяет [1]). Затем останется подобрать постоянные a и a' так, чтобы выполнялись начальные условия (9): для этого достаточно решить систему уравнений

$$\begin{cases} a + a' = 0, \\ aq + a'q' = 1. \end{cases}$$

---

[1]) В самом деле, складывая равенства

$$aq^{n+1} = aq^{n-1} + aq^n \quad \text{и} \quad a'q'^{n+1} = a'q'^{n-1} + a'q'^n,$$

получим равенство

$$aq^{n+1} + a'q'^{n+1} = (aq^{n-1} + a'q'^{n-1}) + (aq^n + a'q'^n).$$

6. Рассмотрим последовательность, отличающуюся от последовательности Фибоначчи тем, что каждый член её равен не сумме, а полусумме двух предыдущих

$$a_{n+2}=\frac{a_n+a_{n+1}}{2} \qquad (n=1,\ 2,\ 3,\ \ldots), \tag{13}$$

при прежних начальных условиях

$$a_1=0, \quad a_2=1.$$

Мы получаем:

$$a_3=\frac{0+1}{2}=\frac{1}{2}, \qquad a_4=\frac{1+\frac{1}{2}}{2}=\frac{3}{4}, \qquad a_5=\frac{\frac{1}{2}+\frac{3}{4}}{2}=\frac{5}{8},$$

$$a_6=\frac{\frac{3}{4}+\frac{5}{8}}{2}=\frac{11}{16}, \qquad a_7=\frac{\frac{5}{8}+\frac{11}{16}}{2}=\frac{21}{32}, \qquad a_8=\frac{\frac{11}{16}+\frac{21}{32}}{2}=\frac{43}{64}, \text{ и т. д.}$$

Общий член a_n возникающей последовательности

$$0,\ 1,\ \frac{1}{2},\ \frac{3}{4},\ \frac{5}{8},\ \frac{11}{16},\ \frac{21}{32},\ \frac{43}{64},\ \ldots \tag{14}$$

может быть найден тем же методом, что и в предыдущем примере; но вместо квадратного уравнения (11) придётся решать уравнение

$$q^2=\frac{1+q}{2},$$

корни которого суть 1 и $\left(-\frac{1}{2}\right)$. Подбирая постоянные a и a' в выражении $a+(-1)^{n-1}\frac{a'}{2^{n-1}}$ так, чтобы удовлетворить начальным условиям (14), мы получаем:

$$a_n=\frac{2}{3}\left\{1+\frac{(-1)^n}{2^{n-1}}\right\} \qquad (n=1,\ 2,\ 3,\ \ldots).$$

7. Если задать последовательность требованием, чтобы каждый её член равнялся среднему арифметическому трёх предыдущих

$$a_{n+3}=\frac{a_n+a_{n+1}+a_{n+2}}{3} \qquad (n=4,\ 5,\ \ldots),$$

то необходимо в качестве начальных условий задать значения трёх первых членов a_1, a_2 и a_3. Вычисление одного за другим всех членов не представляет труда; правда, на пути к построению формулы для общего члена a_n встречаются некоторые затруднения [1]).

Можно указать два различных геометрических представления конечной последовательности чисел

$$a_1,\ a_2,\ a_3,\ \ldots,\ a_n,\ \ldots,\ a_N. \tag{15}$$

---

[1]) Рассуждение, подобное предыдущему, приводит к уравнению третьей степени, имеющему мнимые корни; впрочем, в окончательной формуле мнимости можно избежать.

Первое связано с координатной плоскостью Oxy; в плоскости отмечаются точки с координатами

$$(n,\ a_n) \qquad (n=1,\ 2,\ \ldots,\ N)$$

и затем проводятся вертикальные отрезки, вершины которых находятся в этих точках, а основания — на оси Ox. Члены последовательности изображаются этими вертикальными отрезками («столбиками»). Если член положителен, то отрезок направлен вверх; если отрицателен, то — вниз. Если члены возрастают при возрастании n в некоторых пределах ($n_1 < n < n_2$), то при движении слева направо «столбики» растут (алгебраически); если члены убывают, то «столбики» уменьшаются. При аналитическом задании последовательности

$$a_n = f(n)$$

достаточно наметить график функции $y=f(x)$ и затем прочертить вертикальные отрезки, соответствующие целым абсциссам.

Другое геометрическое представление последовательности чисел связано с одной лишь числовой осью. На числовой оси последовательно отмечаются точки a_1, a_2, ...; чтобы точки не смешались, необходимо ставить при них соответствующие пометки. Если в некоторых пределах $n_1 < n < n_2$ члены последовательности возрастают, то соответствующие точки располагаются слева направо; если убывают, то — справа налево.

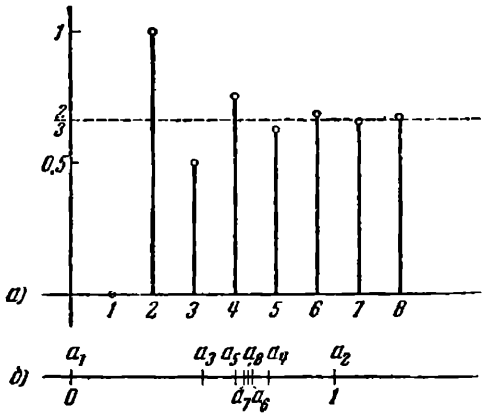


Рис. 67.

Чтобы перейти от первого геометрического представления ко второму, достаточно спроектировать все вертикальные «столбики» на вертикальную же ось и затем (если угодно, придав этой оси горизонтальное положение) поставить пометки при проекциях.

Обоими указанными представлениями пользуются в одинаковой степени и преимущественно с целью дать толчок воображению, не придавая значения собственно-графической стороне дела.

На рис. 67, *а* и *б* дано графическое представление последовательности, определённой рекуррентным соотношением (13), причём значения n взяты только от $n=1$ до $n=8$.

Бесконечные последовательности допускают такие же геометрические представления, что и конечные. Разница лишь в том, что практически выполнить построение, конечно, невозможно и приходится всегда останавливаться на каком-то члене a_n, предоставляя дальнейшее усилиям воображения.

§ 35. Общее определение бесконечной числовой последовательности

В дальнейшем, как это общепринято в математике, говоря о последовательностях, мы будем всегда иметь в виду бесконечные последовательности.

Существенное отличие бесконечной последовательности от конечной заключается в том, что конечная последовательность может быть задана непосредственным перечислением её членов (как бы велико ни было их число), тогда как для бесконечной последовательности такое перечисление принципиально невозможно.

С этим положением легко согласится неискушённый читатель и на вопрос: «чем же в случае бесконечной последовательности можно заменить непосредственное перечисление?» — не замедлит высказаться в том смысле, что бесконечной последовательности должна быть присуща «закономерность», или что должен быть указан некоторый «закон», по которому составляются следующие один за другим члены последовательности; такую же мысль нетрудно найти и во многих литературных источниках.

Оспаривать по существу справедливость такого суждения не приходится; необходимо только подвергнуть более глубокому анализу содержания понятия «закон», или «закономерность», принимая во внимание, что подобного рода термины (способные оказывать сильное, так сказать, гипнотизирующее воздействие) не всегда находят для себя место в лексиконе математика.

Тот, кто спешит произнести слово «закон», по большей части имеет в виду прежде всего математическую формулу, которая выражала бы общий член последовательности в виде функции индекса, или номера, n. Если надлежащим образом уточнить понимание термина «формула», или «элементарная функция» (в том смысле, как это было сделано в § 1), то «указать закон» значит то же самое, что согласно сказанному в § 34 «задать последовательность аналитически», в «замкнутой форме». Этот способ задавать последовательности, конечно, вполне пригоден для бесконечного случая, как и для конечного.

Но подобное толкование «закономерности» слишком ограничено, недостаточно чётко и неустойчиво. Оно, повидимому, не охватывает многих простых «рекуррентных» последовательностей; при этом остаётся открытым вопрос о том, существует или не существует аналитическое представление последовательности в «замкнутой форме».

Во многих случаях подобрать функцию $f(x)$ нетрудно. Так, для последовательности

$$(-1),\ 1,\ (-1),\ 1,\ (-1),\ 1,\ \ldots \tag{16}$$

оказывается возможным положить $f(x)=\cos \pi x$. Более сложным в этом смысле мог бы показаться пример последовательности

$$1,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 0,\ 0,\ \ldots \tag{17}$$

(с чередованием одной единицы и двух нулей); но и в этом случае вопрос, касающийся функции $f(x)$, решается утвердительно: можно положить, например,

$$f(x)=\frac{1}{3}\left\{1+2\cos\left[\frac{2}{3}\pi(x-1)\right]\right\}.$$

Такой же положительный результат можно получить и для произвольной периодической последовательности, с «периодом» p,

$$a_1, a_2, \ldots, a_p, a_1, a_2, \ldots, a_p, a_1, a_2, \ldots, a_p, \ldots, \tag{18}$$

которую можно определить рекуррентным соотношением

$$a_{n+p} = a_n \qquad (n = 1, 2, 3, \ldots)$$

с начальными данными — произвольными значениями $a_1, a_2, a_3, \ldots, a_p$.

С другой стороны, например, рекуррентная зависимость

$$a_{n+1} - a_n = \frac{1}{n+1}$$

с начальным условием $a_1 = 1$, очевидно, определяет единственную последовательность

$$1, 1+\frac{1}{2}, 1+\frac{1}{2}+\frac{1}{3}, \ldots, 1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}, \ldots, \tag{19}$$

которую, казалось бы, нет оснований отнести к числу «незакономерных», хотя вовсе не ясно, существует ли элементарная функция $f(x)$, удовлетворяющая бесконечному числу условий

$$f(n) = 1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n} \qquad (n = 1, 2, 3, \ldots).$$

Вместе с тем нужно обратить внимание и на то обстоятельство, что, опираясь на определение последовательностей «аналитическим» методом, мы ставим понятие последовательности в зависимость от допускаемого класса функций $f(x)$. Ограничиваться только классом «элементарных» функций неестественно. При дальнейшем же расширении понятие «функция» становится столь эластическим аппаратом, что приобретает способность изображать любую эмпирическую закономерность [1]), в связи с чем сама идея «закономерности» поднимается, если можно так выразиться, на более высокую ступень.

В свете высказанных соображений и приведённых примеров не должен показаться удивительным следующий способ определения бесконечной последовательности, который в настоящее время принят в математике.

Числовая последовательность $\{a_n\}$ считается заданной, если указано правило, позволяющее по заданному индексу — произвольному целому положительному числу n — однозначно определить член последовательности a_n, стоящий на n-м месте. Другими словами, *последовательность $\{a_n\}$ задана, если с каждым целым положительным числом n сопоставлено (приведено в соответствие) некоторое число a_n.*

Характер правила, устанавливающего соответствие между n и a_n, безразличен. В простейшем случае «аналитического» определения $a_n = f(n)$ «правило» указывает совокупность элементарных математических операций, которые нужно выполнить над числом n, чтобы получить число a_n. Но «правила» могут быть и совсем иные.

---

[1]) Яркая иллюстрация этой мысли приведена в § 49.

Приведём примеры, в которых «правило» угадывается без труда:

1) 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, ...;

2) 1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, 1, 2, 3, 4, 5, 6, ...;

3) 3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, ...;

4) $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, \ldots$

В каждом из этих примеров «закономерность» (в более широком смысле) имеется налицо: последовательности «заданы», хотя вопрос о существовании функции $f(x)$ крайне неясен, и нельзя также усмотреть наличия простых рекуррентных зависимостей.

В примере 3 правило, определяющее соответствие между n и a_n, таково: a_n есть n-й десятичный знак в разложении отношения длины окружности к длине диаметра.

В примере 4 выписываются в виде последовательности все несократимые правильные рациональные дроби. Порядок определяется тем условием, что из двух дробей $\frac{p}{q}$ и $\frac{p'}{q'}$ дробь $\frac{p'}{q'}$ следует за дробью $\frac{p}{q}$, если $q' > q$; а в случае $q' = q$, если $p' > p$. Чтобы узнать, какая дробь стоит на n-м месте, нужно продолжить последовательность до n-го места и посмотреть, какая дробь станет на этом месте.

Остановимся на некоторых свойствах бесконечных числовых последовательностей.

Среди конечного числа данных различных чисел всегда можно выбрать наибольшее и наименьшее. Понятие наибольшего и наименьшего числа требует небольшого уточнения в тех случаях, когда данные числа не обязательно различны между собой. Рассмотрим сначала конечную последовательность

$$a_1, a_2, a_3, \ldots, a_N. \tag{20}$$

Наибольшим из чисел (20) называется такое, которое не меньше всех остальных: оно обозначается через $\max\{a_1, a_2, \ldots, a_n\}$. Таким образом,

$$\max\{a_1, a_2, \ldots, a_N\} \geqslant a_n \qquad (n = 1, 2, \ldots, N), \tag{21}$$

но вместе с тем равенство имеет место хотя бы для одного значения n.

Точно так же *наименьшим* из чисел (20) называется такое, которое не больше всех остальных: оно обозначается через $\min\{a_1, a_2, \ldots, a_n\}$; мы получаем

$$\min\{a_1, a_2, \ldots, a_N\} \leqslant a_n \qquad (n = 1, 2, \ldots, N) \tag{22}$$

при наличии равенства хотя бы для одного значения n.

Существование наибольшего и наименьшего членов конечной числовой последовательности не вызывает никаких сомнений [1]).

Совсем иначе обстоит дело в случае бесконечной числовой последовательности: такая последовательность *может не иметь наибольшего члена и может не иметь наименьшего.*

Так, последовательность натуральных чисел $\{n\}$ не имеет наибольшего члена; последовательность целых отрицательных чисел

[1]) Не исключается такой случай, что наибольший и наименьший члены совпадают.

$\{-n\}$ не имеет наименьшего; последовательность, составленная из целых чисел, $\{(-1)^{n+1}n\}$ не имеет ни наибольшего, ни наименьшего.

Другое яркое отличие бесконечных числовых последовательностей от конечных связано с понятием *ограниченности*.

Последовательность (всё равно какая — конечная или бесконечная) называется

а) *ограниченной сверху* (говорят также — *справа*), если все члены её меньше одного и того же числа M:

$$a_n < M \qquad (n = 1, 2, 3, \ldots), \tag{23}$$

b) *ограниченной снизу* (*слева*), если все члены её больше одного и того же числа m:

$$a_n > m \qquad (n = 1, 2, 3, \ldots), \tag{24}$$

c) *ограниченной* (*просто*), если она ограничена и сверху и снизу:

$$m < a_n < M \qquad (n = 1, 2, 3, \ldots). \tag{25}$$

Всякая конечная последовательность, очевидно, ограничена. В качестве M можно взять любое число, большее чем наибольший из членов последовательности; в качестве m — любое число, меньшее чем наименьший из членов последовательности.

Аналогичное утверждение неверно для бесконечных последовательностей. Примером бесконечной последовательности, не ограниченной сверху, может служить последовательность натуральных чисел $\{n\}$; но она ограничена снизу. Напротив, последовательность $\{-n\}$ не ограничена снизу. Последовательность $\{(-1)^{n+1}n\}$ не ограничена ни сверху, ни снизу.

Бесконечные арифметические прогрессии первого порядка ограничены снизу, но не сверху, или наоборот, смотря по тому, будет ли их разность положительной или отрицательной. Бесконечные геометрические прогрессии ограничены снизу, но не сверху, если их знаменатель больше единицы; ограничены и сверху и снизу, если знаменатель заключён между -1 и $+1$; не ограничены ни сверху, ни снизу, если знаменатель меньше чем -1. Бесконечная последовательность Фибоначчи (10) не ограничена сверху, но последовательность (14) ограничена. Ограничены сверху и снизу все периодические последовательности, а также последовательности 1, 3 и 4 на стр. 151; но последовательность 2 (там же) ограничена снизу, но не сверху.

Последовательность

$$\left\{1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right\} \tag{19}$$

не ограничена сверху. Чтобы в этом убедиться, нужно установить, что, как бы велико ни было число M, всегда можно указать такой

член последовательности a_n, что будет иметь место неравенство $a_n \geqslant M$; нужно, другими словами, указать такое n, что

$$1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}\geqslant M.$$

От нас зависит взять n вида $n=2^p$. Обратим внимание на то, что

$$1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^p}=$$

$$=\left\{\begin{array}{l} 1+ \\ +\frac{1}{2}+ \\ +\frac{1}{3}+\frac{1}{4}+ \\ +\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+ \\ +\ldots\ldots\ldots+ \\ +\frac{1}{2^{p-1}+1}+\ldots+\frac{1}{2^p} \end{array}\right\} > \left\{\begin{array}{l} 1+ \\ +\frac{1}{2}+ \\ +\frac{1}{4}+\frac{1}{4}+ \\ +\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+ \\ +\ldots\ldots\ldots+ \\ +\frac{1}{2^p}+\ldots+\frac{1}{2^p} \end{array}\right\} > \frac{p}{2}\ ^{1}).$$

Поэтому, взяв $p>2M$, будем иметь:

$$1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^p}>M.$$

Если бесконечная последовательность ограничена сверху, то отсюда не следует, что в ней имеется наибольший член. Так, для последовательности

$$\left\{\frac{n}{n+1}\right\}\equiv\frac{1}{2},\ \frac{2}{3},\ \frac{3}{4},\ldots$$

общий член $\frac{n}{n+1}$ меньше чем $1\ (\equiv M)$; но в ней нет наибольшего члена.

Обратное, напротив, верно: совершенно очевидно, что если в последовательности $\{a_n\}$ имеется наибольший член, то, взяв M хоть немногим больше этого члена, мы получим для всех членов последовательности неравенство

$$a_n<M.$$

§ 36. Теорема Больцано-Вейерштрасса о существовании предельной точки

Многие из терминов, которыми мы будем пользоваться дальше, встречались неоднократно как в этой статье, так и в предшествующих. Уточним, однако, какой смысл мы будем в них вкладывать в настоящем параграфе.

---

[1]) В самом деле, сумма в каждой строчке $\geqslant\frac{1}{2}$, а всего строчек $p+1$.

Каждое действительное число x мы представляем себе написанным в виде бесконечной десятичной дроби

$$x = x_0, x_1 x_2 x_3 \ldots,$$

где x_0 обозначает число единиц, x_1 — число десятых, x_2 — число сотых, ..., вообще x_n — число единиц разряда 10^{-n}, заключающихся в разложении числа x.

Числа, представимые в виде конечной десятичной дроби, мы условимся записывать также в виде бесконечной десятичной дроби — с «нулём в периоде», например

$$4{,}83 = 4{,}83000 \ldots$$

Такие числа допускают ещё другое представление, именно, с «девяткой в периоде», например

$$4{,}83 = 4{,}82999 \ldots$$

Итак, каждое число может быть представлено в виде бесконечной десятичной дроби одним или — самое большее — двумя различными способами.

Но, с другой стороны, всякая бесконечная десятичная дробь представляет л и ш ь о д н о число.

Под «десятичными дробями нулевого ранга» мы будем понимать все целые числа и обозначать их совокупность через (D_0); под «десятичными дробями первого ранга» будем понимать все конечные десятичные дроби, имеющие не более одной значащей цифры после запятой, и обозначать их совокупность через (D_1); вообще под десятичными дробями «n-го ранга» будем понимать все конечные десятичные дроби, имеющие не более n значащих цифр после запятой, и будем обозначать их совокупность через (D_n) и т. д. Каждая из числовых совокупностей (D_n), будучи расположена в порядке возрастания, представляет собой арифметическую прогрессию (простирающуюся в о б е стороны — вправо и в л е в о), с разностью, равной 10^{-n}.

Под *промежутком* $[a, b]$, где $a < b$, условимся понимать совокупность всех действительных чисел x, заключённых между a и b, в к л ю ч а я эти две точки (начальную и конечную):

$$a \leqslant x \leqslant b.$$

Точки промежутка, отличные от начальной и конечной, называются *внутренними*; начальная и конечная точки называются также *концевыми*. *Длиной* промежутка $[a, b]$ называется разность $b - a$; длина промежутка — всегда положительное число. *Центром* промежутка $[a, b]$ называется число (точка) $\frac{a+b}{2}$.

Если a и b — две произвольные точки ($a \leqslant b$), то они образуют или не образуют промежуток, смотря по тому, будет ли $a \neq b$, или $a = b$. В первом случае начальной точкой промежутка, образован-

ного числами a и b, является меньшее из этих чисел, конечной — большее. Длина промежутка, образованного точками a и b, равна $|a-b|$. Выражение $|a-b|$ называется также *отклонением*, или *расстоянием*, точки a от точки b (или точки b от точки a); очевидно, при любых a и b

$$|a-b| \geqslant 0. \tag{26}$$

По определению абсолютной величины неравенство

$$|x-c| < m \qquad (m > 0) \tag{27}$$

означает совместное выполнение двух неравенств

$$\begin{cases} x-c < m, \\ -(x-c) < m, \end{cases}$$

что равносильно выполнению одного двойного неравенства

$$c-m < x < c+m. \tag{28}$$

Таким образом, из неравенства вида (27) следует неравенство (28), и обратно, из неравенства вида (28) следует неравенство (27). Так как точка c есть центр промежутка $[c-m,\ c+m]$, а число $2m$ — его длина, то сказанное можно выразить ещё так: если отклонение точки x от точки c меньше чем m, то точка x расположена внутри промежутка длины $2m$ с центром c, и обратно.

Окрестностью точки c называется совокупность внутренних точек любого промежутка, для которого c является внутренней точкой. Чаще всего приходится иметь дело с окрестностями, для которых данная точка является центром. Под *m-окрестностью* точки c $(m > 0)$ понимают совокупность точек, расстояние которых от точки c меньше чем m; чтобы точка x принадлежала m-окрестности точки c, необходимо и достаточно выполнение любого из соотношений (27) или (28). Если $0 < m_1 < m_2$ и точка x принадлежит m_1-окрестности точки c, то она принадлежит также и её m_2-окрестности; обратное неверно.

Читателю рекомендуется ясно представлять себе геометрический смысл всего предыдущего (на числовой оси).

Отметим ещё в качестве леммы следующее положение:

Если имеется «двусторонняя» [1] *арифметическая прогрессия с разностью d (>0), то при условии $m > d$ m-окрестность любой точки c, не принадлежащей прогрессии, содержит обе точки прогрессии, между которыми лежит c, а следовательно, и весь образованный ими промежуток. Если же сама точка c входит в состав прогрессии, то её m-окрестность содержит обе ближайшие точки прогрессии, а также оба прилежащих промежутка длины d.*

---

[1]) То-есть продолжающаяся неограниченно не только вправо, но и влево.

Это совершенно ясно геометрически. Аналитическое доказательство также несложно. Пусть точки прогрессии имеют вид

$$a+(p-1)\,d, \qquad (-\infty<p<+\infty)$$

и пусть

$$a+(p-1)\,d\leqslant c\leqslant a+pd;$$

тогда с помощью условия $m>d$ получаем

$$c-m\leqslant a+pd-m<a+pd-d=a+(p-1)\,d,$$

$$c+m\geqslant a+(p-1)\,d+m>a+(p-1)\,d+d=a+pd,$$

так что

$$c-m<a+(p-1)\,d<a+pd<c+m.$$

Теперь мы сформулируем основную теорему этого параграфа.

Какова бы ни была ограниченная бесконечная последовательность

$$\{a_n\}=a_1,\ a_2,\ a_3,\ \ldots,\ a_n,\ \ldots, \tag{29}$$

всегда существует по меньшей мере одна точка c, обладающая тем свойством, что в любой её сколь угодно малой (по длине) *окрестности имеется бесконечное множество точек последовательности* (29).

Доказательство этой теоремы базируется на следующем очевидном принципе, который мы формулируем для большей наглядности следующим образом: *если в конечном числе ящиков разложено бесконечное множество предметов, то хотя бы в одном ящике окажется бесконечное множество этих предметов.*

Обращаясь к доказательству теоремы, условимся прежде всего, что если среди чисел a_n будут встречаться отрицательные, то мы будем записывать их в виде десятичных дробей с отрицательными характеристиками, но положительными мантиссами, например

$$-4{,}52038\ \ldots=\bar{5}{,}47961\ \ldots$$

Все члены последовательности разбиваются на классы («распределяются по ящикам»), смотря по характеристике, т. е. по тому, какое число целых единиц стоит в соответствующей записи слева от запятой. Число таких классов — конечное, так как последовательность $\{a_n\}$ — ограниченная, и потому все числа a_n заключены между двумя какими-то числами m и M:

$$m<a_n<M \qquad (n=1,\ 2,\ 3,\ \ldots).$$

Не ограничивая общности, можно считать числа m и M целыми. Наших классов, или «ящиков», — столько же, сколько различных характеристик

$$m,\ m+1,\ m+2,\ \ldots,\ M-1, \tag{30}$$

именно, $N=M-m$. Так как членов последовательности $\{a_n\}$ бесконечное множество, то хотя бы в одном из «ящиков» их окажется

бесконечное множество. Другими словами, среди целых чисел (30) найдётся одно такое (обозначим его через c_0), что бесконечное множество членов последовательности $\{a_n\}$ имеют характеристику c_0, т. е. удовлетворяют неравенству

$$c_0 \leqslant x < c_0 + 1.$$

Предположим, далее, что «ящик» с характеристикой c_0 разбивается на 10 «подразделений» — смотря по тому, какова первая цифра после запятой в рассматриваемой десятичной дроби. Так как в этот «ящик» попадает бесконечное множество членов последовательности, то хотя бы в одно из 10 «подразделений» их также попадает бесконечное множество. Пусть, например, бесконечное множество членов последовательности попадает в «подразделение», характеризуемое цифрой c_1, на первом месте справа от запятой; другими словами, допустим, что имеется бесконечное множество чисел a_n, удовлетворяющих неравенству

$$c_0,c_1 \leqslant x < c_0,c_1 + \frac{1}{10}.$$

Вообразим ещё дальше, что «подразделение», характеризуемое цифрами c_0,c_1, разбивается ещё на 10 «под-подразделений» — в зависимости от второй цифры справа от запятой. Хотя бы одно из таких под-подразделений, например характеризуемое цифрой c_2 на втором месте справа от запятой, содержит бесконечное множество членов последовательности, так что существует бесконечное множество членов, удовлетворяющих неравенству

$$c_0,c_1c_2 \leqslant x < c_0,c_1c_2 + \frac{1}{100}.$$

Можно представить себе, что подобного рода процесс продолжается до бесконечности. Мы получим таким образом бесконечную последовательность ц и ф р [1])

$$c_0, c_1, c_2, \ldots, c_n, \ldots,$$

соответственно стоящих слева от запятой, на первом, на втором, ..., на n-м месте справа от запятой, ..., и обладающих тем свойством, что в каждом из промежутков

$$\left.\begin{array}{c} [c_0;\ c_0+1], \\ [c_0,c_1;\ c_0,c_1+10^{-1}], \\ [c_0,c_1c_2;\ c_0,c_1c_2+10^{-2}], \\ \ldots\ldots\ldots\ldots \\ [c_0,c_1c_2\ldots c_n;\ c_0,c_1c_2\ldots c_n+10^{-n}], \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right\} \qquad (31)$$

содержится бесконечное множество членов последовательности $\{a_n\}$.

[1]) Впрочем, c_0 может быть каким угодно, положительным или отрицательным, целым числом или нулём.

Рассмотрим теперь число c, выражающееся бесконечной дробью

$$c = c_0,c_1c_2\ldots c_n\ldots, \tag{32}$$

и докажем, что в любой его окрестности содержится бесконечное множество точек последовательности $\{a_n\}$. Это число принадлежит одновременно каждому из промежутков (31). Возьмём какую угодно окрестность точки c, т. е. сколь угодно малый промежуток, содержащий внутри точку c. Не ограничивая общности, можно допустить, что точка c есть центр этого промежутка [1]). Обозначая длину промежутка через 2ε, мы будем иметь дело с ε-окрестностью точки c:

$$|x - c| < \varepsilon,$$

или

$$c - \varepsilon < x < c + \varepsilon.$$

Десятичные дроби ранга n образуют арифметическую прогрессию с разностью 10^{-n}. Если только

$$\varepsilon > 10^{-n}, \tag{33}$$

то по лемме ε-окрестность точки c содержит обе ближайшие к c дроби

$$c_0,c_1c_2\ldots c_n \quad \text{и} \quad c_0,c_1c_2\ldots c_n + 10^{-n},$$

а следовательно, содержит и образованный ими промежуток

$$(c_0,c_1c_2\ldots c_n;\ \ c_0,c_1c_2\ldots c_n + 10^{-n}) \tag{34}$$

(это верно также и в том случае, если c совпадает с одной из названных дробей). Но так как в промежутке (34) содержится бесконечное множество членов последовательности $\{a_n\}$, то справедливо то же самое и относительно рассматриваемой ε-окрестности точки c.

Остаётся заметить, что неравенство (33) равносильно следующему:

$$n > \lg \frac{1}{\varepsilon}. \tag{35}$$

Итак, для того чтобы предыдущее рассуждение имело силу, достаточно, пользуясь десятичными дробями ранга n, взять в качестве n любое целое число, удовлетворяющее неравенству (35).

Теорема доказана.

Всякая точка, обладающая тем свойством, что в любой её окрестности содержится бесконечное множество точек последовательности $\{a_n\}$, называется *предельной точкой* этой последовательности.

Пользуясь этим термином, можно сформулировать теорему Больцано-Вейерштрасса следующим образом:

---

[1]) Если доказано, что в промежутке AA' (с центром c) содержится бесконечное множество точек последовательности $\{a_n\}$, то тем самым аналогичное утверждение будет доказано и для промежутка AB, охватывающего AA'.

Всякая ограниченная бесконечная последовательность $\{a_n\}$ имеет по крайней мере одну предельную точку.

Примечание 1. При доказательстве теоремы Больцано-Вейерштрасса то обстоятельство, что множество рассматриваемых чисел a_n ($n = 1, 2, 3, \ldots$) может быть расположено в порядке последовательности, т. е. что оно — счётное (см. Э. э. м., т. I, стр. 92), не играет никакой роли. Поэтому теорема справедлива не только для последовательности чисел, но и для любого бесконечного множества чисел E, заключённого в данном конечном промежутке.

Примечание 2. Определение предельной точки множества E можно видоизменить следующим образом: точка c называется предельной точкой данного множества E, если в любой её сколь угодно малой окрестности имеется хоть одна точка этого множества, отличная от точки c.

В самом деле, если верно, что в любой окрестности точки c содержится хоть одна точка данного множества E, то верно и то, что в любой окрестности их содержится сколько угодно. Пусть, например, $(c-\varepsilon,\ c+\varepsilon)$ есть данная ε-окрестность точки c. Выберем в ней точку c_1 данного множества E, отличную от c, затем возьмём ε_1-окрестность точки c, подчиняя ε_1 условию $\varepsilon_1 < |c_1 - c|$, и в этой ε_1-окрестности $(c-\varepsilon_1,\ c+\varepsilon_1)$ выберем новую точку c_2 множества E, отличную от c; затем возьмём ε_2-окрестность той же точки c, подчиняя ε_2 условию $\varepsilon_2 < |c_2 - c|$, и в этой ε_2-окрестности $(c-\varepsilon_2,\ c+\varepsilon_2)$ выберем третью точку c_3 данного множества E, отличную от c, и т. д.

§ 37. Примеры. Предел как единственная предельная точка

Различные последовательности могут иметь то или иное число предельных точек; существуют также последовательности, обладающие бесконечным множеством предельных точек. Заметим ещё, что предельная точка может «принадлежать» последовательности, т. е. входить в состав её членов (может быть, даже неоднократно), или

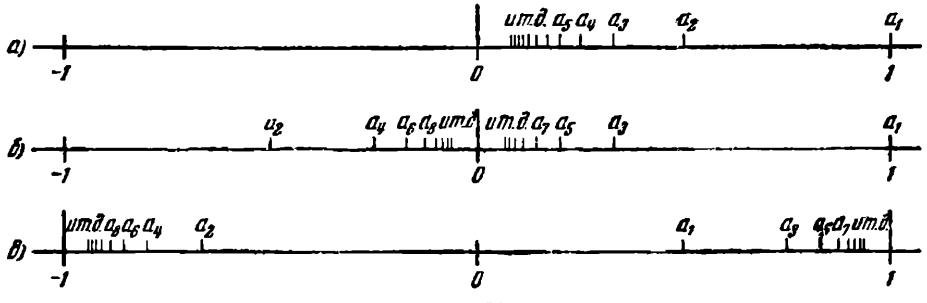


Рис. 68.

не принадлежать ей. Разнообразие мыслимых возможностей видно из приведённых ниже примеров. Читателю рекомендуется при рассмотрении примеров пользоваться геометрическим представлением, изображая члены последовательности в виде точек на числовой оси.

1. Последовательность

$$\left\{\frac{1}{n}\right\} \equiv 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

имеет одну предельную точку 0 (рис. 68, *а*).

2. То же можно сказать о последовательности

$$0,\ 1,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4},\ \ldots,$$

которая от предыдущей отличается тем, что на первое место поставлен нуль и все члены сдвинуты на одно место; также о последовательности

$$0,\ 1,\ 0,\ \frac{1}{2},\ 0,\ \frac{1}{3},\ 0,\ \frac{1}{4},\ \ldots,$$

которая отличается от предыдущей тем, что нуль вставлен между всякими двумя членами (начиная со второго).

3. То же можно сказать о последовательностях вида

$$\left\{\frac{1}{n^{\alpha}}\right\} \equiv 1,\ \frac{1}{2^{\alpha}},\ \frac{1}{3^{\alpha}},\ \frac{1}{4^{\alpha}},\ \ldots$$

при условии $\alpha > 0$.

4. То же можно сказать о последовательности с чередующимися знаками (рис. 68, *б*)

$$\left\{\frac{(-1)^{n+1}}{n}\right\} \equiv 1,\ -\frac{1}{2},\ \frac{1}{3},\ -\frac{1}{4},\ \ldots$$

и о любой последовательности, возникающей из последовательности $\left\{\frac{1}{n}\right\}$ посредством расстановки знаков совершенно произвольным образом.

5. Любая бесконечная геометрическая прогрессия имеет единственную предельную точку 0, если только её знаменатель q удовлетворяет условию $|q| < 1$.

6. Последовательности

$$\left\{a+\frac{1}{n}\right\};\ \left\{a+\frac{1}{n^{\alpha}}\right\},\ \text{где}\ \alpha>0;\ \left\{a+\frac{(-1)^n}{n}\right\};$$
$$\{a+q^n\},\ \text{где}\ |q|<1,$$

имеют единственную предельную точку a.

7. Последовательность

$$\left\{(-1)^{n+1}\frac{n}{n+1}\right\} \equiv \frac{1}{2},\ -\frac{2}{3},\ \frac{3}{4},\ -\frac{4}{5},\ \ldots$$

имеет д в е предельные точки: 1 и -1 (рис. 68, *в*).

8. То же можно сказать о последовательности

$$\{\cos n\pi\} \equiv \{(-1)^n\} \equiv -1,\ 1,\ -1,\ 1,\ \ldots$$

9. Если последовательность $\{a_n\}$ имеет единственную предельную точку a, а последовательность $\{b_n\}$ — единственную предельную точку b, причём $b \neq a$, то последовательность

$$a_1,\ b_1,\ a_2,\ b_2,\ a_3,\ b_3,\ \ldots \qquad (36)$$

имеет две предельные точки: a и b. Если же $b=a$, то предыдущая последовательность имеет лишь одну предельную точку, именно, $a=b$, например

$$\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{2}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{1}{5}, \frac{2}{5}, \ldots$$

Ничто существенно не изменилось бы в этом примере, если бы в рассматриваемой последовательности (36), «составленной» из последовательностей $\{a_n\}$ и $\{b_n\}$, члены этих последовательностей не чередовались между собой, а следовали бы в каком угодно порядке.

10. Последовательность

$$\left\{\sin n\frac{\pi}{2}\right\} \equiv 1,\ 0,\ -1,\ 0,\ 1,\ 0,\ -1,\ 0, \ldots$$

имеет три предельные точки: 1, -1 и 0. Любая периодическая последовательность имеет в качестве точек все те числа, которые встречаются в пределах периода.

11. Если последовательность $\{a_n\}$ имеет единственную предельную точку a, последовательность $\{b_n\}$ — единственную предельную точку $b, \ldots$, последовательность $\{l_n\}$ — единственную предельную точку l, то при условии, что числа $a, b, \ldots, l$ различны между собой, все они являются предельными точками последовательности

$$a_1, b_1, \ldots, l_1, \quad a_2, b_2, \ldots, l_2, \quad a_3, b_3, \ldots, l_3, \ldots$$

В случае совпадений число предельных точек соответствующим образом уменьшается.

12. Если рассмотреть последовательности, стоящие в первой, второй, третьей и т. д. строчках следующей таблицы:

$$\begin{array}{llll}
\frac{1}{1}+\frac{1}{1^2}, & \frac{1}{1}+\frac{1}{2^2}, & \frac{1}{1}+\frac{1}{3^2}, & \frac{1}{1}+\frac{1}{4^2}, \ldots, \\
\frac{1}{2}+\frac{1}{1^2}, \nearrow & \frac{1}{2}+\frac{1}{2^2}, \nearrow & \frac{1}{2}+\frac{1}{3^2}, \nearrow & \frac{1}{2}+\frac{1}{4^2}, \ldots, \\
\frac{1}{3}+\frac{1}{1^2}, \nearrow & \frac{1}{3}+\frac{1}{2^2}, \nearrow & \frac{1}{3}+\frac{1}{3^2}, \nearrow & \frac{1}{3}+\frac{1}{4^2}, \ldots, \\
\frac{1}{4}+\frac{1}{1^2}, \nearrow & \frac{1}{4}+\frac{1}{2^2}, \nearrow & \frac{1}{4}+\frac{1}{3^2}, \nearrow & \frac{1}{4}+\frac{1}{4^2}, \ldots, \\
\ldots & \ldots & \ldots & \ldots
\end{array}$$

то легко убедиться, что каждая из них имеет единственную предельную точку, а именно, 1, $\frac{1}{2}$, $\frac{1}{3}$ и т. д. Но составляя последовательность из всех элементов таблицы по диагоналям (как показывают стрелочки), получаем последовательность

$$\frac{1}{2}, \frac{3}{2}, \frac{5}{4}, \frac{4}{3}, \frac{3}{4}, \frac{10}{9}, \frac{5}{4}, \frac{7}{12}, \frac{11}{18}, \frac{17}{16}, \ldots,$$

для которой предельными точками являются все числа

$$1, \frac{1}{2}, \frac{1}{3}, \text{ и т. д.}$$

и ещё число 0 (см. по этому поводу теорему на стр. 163).

13. Последовательность 1 на стр. 151 имеет две предельные точки: 0 и 1.

14. Последовательность 3 (там же) не может иметь иных предельных точек, кроме десяти следующих:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9.$$

Все ли они являются предельными точками этой последовательности, — в настоящее время нельзя считать установленным.

15. Последовательность 4 (там же) имеет в качестве предельных точек все числа x без исключения из промежутка $[0, 1]$:

$$0 \leqslant x \leqslant 1.$$

Действительно, в любом промежутке содержится бесконечное множество рациональных чисел (рациональные числа расположены всюду плотно [1]).

16. Последовательность $\{\sin n\pi\theta\}$, где θ — иррациональное число, имеет предельными точками все точки промежутка $[-1, +1]$:

$$-1 \leqslant x \leqslant +1.$$

Доказательство теоремы проведём «от противного».

Вообразим, что на единичной окружности отмечены, как принято в тригонометрии, все точки P_n, соответствующие центральным углам вида $n\pi\theta$, а также на вертикальном диаметре — их проекции P'_n ($n = 1, 2, 3, \ldots$). Если бы на этом диаметре существовал хоть один промежуток, на котором бы не было точек P'_n, то на окружности, очевидно, нашёлся бы «дуговой» промежуток, на котором бы не было точек P_n. Таких промежутков на окружности может быть больше чем один, но так как сумма их длин не превышает длины окружности 2π, то среди них можно выделить наибольший (или один из наибольших, если их несколько). Обозначим через L дугу окружности, соответствующую этому промежутку, и через λ — её длину ($0 < \lambda < 2\pi$). Вращая дугу L около центра на угол $2n\pi$, кратный 2π, мы получаем новую дугу L_n, однако соответствующую тому же самому дуговому промежутку.

Условимся теперь называть *гомологическими* между собой такие дуги, которые получаются одна из другой посредством вращения на угол, кратный $\pi\theta$.

Если дуга L' — гомологическая по отношению к L, то ясно, что на ней также нет точек P_n.

Рассмотрим совокупность дуг L_n ($n = 1, 2, 3, \ldots$), гомологических дуге $L \equiv L_0$. Так как длина каждой из них равна $\lambda > 0$, а число их — бесконечное, то наверное какие-нибудь две из них, например L_p и L_q ($p < q$), окажутся покрывающими на окружности один и тот же дуговой промежуток M длины $\mu\ (> 0)$.

---

[1]) См. Э. э. м., кн. 1, А. Я. Хинчин, Элементы теории чисел, гл. IV и V.

Возможны два случая: 1) $\mu=\lambda$ и 2) $\mu<\lambda$.

Первый случай приходится отвергнуть. В самом деле, если $\mu=\lambda$, то у дуг L_p и L_q совпадают как начальные, так и конечные точки; значит, дуга L_q получается из дуги L_p в результате вращения на угол, который равен $(q-p)\pi\theta$; с другой же стороны, этот угол является кратным 2π. В таком случае

$$(q-p)\pi\theta=2r\pi \qquad (r\text{ — целое}),$$

т. е.

$$\theta=\frac{2r}{q-p},$$

что противоречит предположению относительно иррациональности θ.

Но приходится отвергнуть и второй случай. Действительно, в дуговом промежутке, покрываемом объединением дуг L_p и L_q, не содержится точек P_n, так как их нет ни на дуге L_p, ни на дуге L_q; длина же его равна

$$2\lambda-\mu>\lambda,$$

что противоречит определению дуги L.

Теорема. *Если некоторое множество точек E имеет в данном* (конечном) *промежутке бесконечное множество предельных точек, то всякая предельная точка множества, составленного из этих предельных точек, есть вместе с тем предельная точка данного множества.*

Доказательство. Пусть E' есть множество предельных точек данного множества E и пусть c'' есть предельная точка множества E'.

По предположению, в любой окрестности точки c'' имеется сколько угодно точек множества E', и вместе с тем в любой окрестности каждой из точек c' множества E' имеется сколько угодно точек множества E; требуется доказать, что в любой окрестности точки c'' имеется хоть одна[1]) точка данного множества E, отличная от c''.

Пусть $(c''-\varepsilon,\ c''+\varepsilon)$ — заданная окрестность точки c''. В окрестности $\left(c''-\frac{\varepsilon}{2},\ c''+\frac{\varepsilon}{2}\right)$ этой точки мы найдём точку c' множества E' и будем иметь:

$$|c'-c''|<\frac{\varepsilon}{2}.$$

В окрестности $\left(c'-\frac{\varepsilon}{2},\ c'+\frac{\varepsilon}{2}\right)$ точки c' существует сколько угодно точек данного множества E; выберем одну из них, например c, отличную от c'', так что будет:

$$|c-c'|<\frac{\varepsilon}{2}.$$

Тогда по известному свойству абсолютных величин[2]) получим:

$$|c-c''|\leqslant|c-c'|+|c'-c''|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

что и требовалось доказать.

Примечание. Теорема остаётся справедливой, и сохраняется в силе доказательство, если «множество E» заменить «последовательностью $\{a_n\}$».

---

[1]) См. примечание 2 на стр. 159.

[2]) См. сноску [2]) на стр. 170.

Если ограниченная последовательность имеет только одну предельную точку, то эта точка называется пределом последовательности.

Если же ограниченная последовательность имеет более одной предельной точки, то говорят, что последовательность не имеет предела.

Таким образом, каждая из последовательностей, приведённых в примерах 1—6, имеет предел, а последовательности в примерах 7—16 предела не имеют [1]).

Если точка c есть не только предельная точка, но и предел ограниченной последовательности $\{a_n\}$, то верно не только то, что в любой окрестности c лежит бесконечное множество точек последовательности, но и то, что в любой окрестности c лежат почти все точки последовательности. При этом «почти все» означает—«все, кроме, может быть, конечного числа».

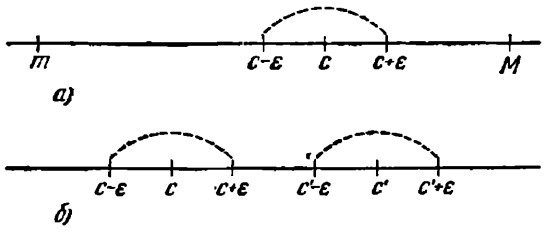


Рис. 69.

Желая это установить, достаточно, как и раньше, ограничиться рассмотрением промежутков, по отношению к которым c является центром. Итак, пусть дана некоторая ε-окрестность точки c (рис. 69,*а*); считая, что все точки последовательности $\{a_n\}$ заключены между m и M, можно допустить, что $\varepsilon < M - c$ и $\varepsilon < c - m$, так что промежуток $[c-\varepsilon, c+\varepsilon]$ лежит внутри промежутка (m, M). Будем рассуждать «от противного»: все точки последовательности $\{a_n\}$ лежат в промежутке $[m, M]$, и если бы было неверно, что «почти все» они лежат в промежутке $[c-\varepsilon, c+\varepsilon]$, то это означало бы, что бесконечное множество их заключается в паре промежутков $[m, c-\varepsilon]$ и $[c+\varepsilon, M]$, и следовательно, заключается хотя бы в одном из них, например в промежутке $[m, c-\varepsilon]$. Но по теореме Больцано-Вейерштрасса отсюда следовало бы, что последовательность $\{a_n\}$ имеет предельную точку в промежутке $(m, c-\varepsilon)$. Последнее невозможно, так как, по предположению, c — единственная предельная точка последовательности. Мы приходим, таким образом, к заключению, что «почти все» точки последовательности лежат в промежутке $[c-\varepsilon, c+\varepsilon]$.

Обратно, если верно, что в любой окрестности точки c лежат «почти все» точки бесконечной последовательности $\{a_n\}$, то отсюда

---

[1]) Это относится и к примеру 14. Можно утверждать, что последовательность, приведённая в этом примере, имеет по крайней мере две предельные точки. Иначе число π представлялось бы периодической десятичной дробью и в таком случае было бы рациональным (см. Э. э. м., кн. 1, стр. 312, теорема 4), тогда как доказано, что π — число иррациональное.

следует, что c есть единственная предельная точка этой последовательности. Докажем и это «от противного». Пусть, кроме точки c, есть ещё предельная точка c' ($\neq c$); предположим, например, что $c' > c$ (рис. 69, *б*). Так как c' — предельная точка, то в любой её ε-окрестности содержится бесконечное множество точек последовательности. Что касается точки c, то в её ε-окрестности содержатся «почти все» точки последовательности. Выбрав ε удовлетворяющим неравенству $\varepsilon < \frac{c'-c}{2}$, получим $c+\varepsilon < c'-\varepsilon$.

Две ε-окрестности оказались не имеющими ни одной общей точки; мы пришли к противоречию.

Подвергнем, наконец, перефразировке выражение «почти все» точки последовательности. Если «почти все» точки последовательности заключены в некоторой ε-окрестности точки c, то это значит, что лишь конечное их число находится вне этого промежутка. Каждый из этих членов последовательности имеет свой индекс, и раз число их конечно, то среди этих индексов есть наибольший. Обозначим этот наибольший индекс через N. Он зависит, конечно, от выбранной окрестности точки c, т. е. зависит от числа ε; чтобы подчеркнуть это обстоятельство, иногда при N ставят значок ε и вместо N пишут N_ε. Итак, среди членов последовательности $\{a_n\}$, у которых индекс n не превышает N, одни, может быть, попадают в ε-окрестность точки c, другие оказываются вне её; но если только $n > N$, то можно уже наверное сказать, что соответствующий член a_n находится в ε-окрестности.

Высказанные соображения позволяют иначе сформулировать понятие предела числовой последовательности. Это иное определение, к которому мы сейчас обратимся (исторически раньше сложившееся [1]), логически строго равносильно приведённому выше.

§ 38. Предел последовательности: классическое определение и основные свойства

Число a [2] *называется пределом* (конечным пределом, пределом «в собственном смысле» — см. § 39) *числовой последовательности*

$$\{a_n\} = a_1, a_2, a_3, \ldots, a_n, \ldots,$$

если, как бы мало ни было наперёд заданное положительное число ε, *можно указать такое число* $N \equiv N_\varepsilon$, *что неравенство*

$$n > N \tag{37}$$

влечёт за собой неравенство

$$|a_n - a| < \varepsilon. \tag{38}$$

[1]) О. Коши впервые дал то определенне предела, которое в настоящее время является общеизвестным и повседневно употребляемым.

[2]) Для удобства дальнейшего изложения буква c предыдущего параграфа заменена здесь буквой a.

Или словами (что гораздо менее отчётливо): *число a есть предел последовательности $\{a_n\}$, если отклонение общего члена последовательности от a делается сколь угодно малым при условии, что индекс члена достаточно велик.*

Для обозначения того факта, что a есть предел последовательности $\{a_n\}$, применяется любая из записей:

1) $\lim a_n = a$ (более старая),

2) $a_n \to a$ (более новая).

В наше время обе они одинаково употребительны.

Вместо «a есть предел a_n» пользуются постоянно синонимическим оборотом: «a_n стремится к a». Если хотят указать, что последовательность имеет предел, не указывая, — какой именно, то говорят, что последовательность — *сходящаяся*. Иногда говорят также «последовательность сходится к a» или «к пределу a».

Если последовательность не имеет предела, её называют *расходящейся*.

Рассмотрим теперь внимательнее примеры 1—5 § 37 с точки зрения этого определения. Во всех этих примерах предел a равен нулю, и потому неравенство (38), наличие которого требуется установить, принимает более простой вид

$$|a_n| < \varepsilon. \tag{39}$$

В примере 1 $a_n = \frac{1}{n}$, и неравенство (39) выполняется при $n > \frac{1}{\varepsilon}$, так что в качестве N может быть взято любое число, не меньшее чем $\frac{1}{\varepsilon}$.

В первой последовательности примера 2 $a_n = \frac{1}{n-1}$ (при $n \geqslant 2$) и неравенство (39) выполняется, если только $n > \frac{1}{\varepsilon} + 1$; следовательно, в качестве N может быть взято любое число, хотя бы на единицу большее чем $\frac{1}{\varepsilon}$.

Во второй последовательности примера 2 $a_{2m} = \frac{1}{m}$, $a_{2m-1} = 0$, и неравенство (39) выполняется автоматически для всех нечётных n; что касается чётных, то оно выполнено, если $\frac{n}{2} = m > \frac{1}{\varepsilon}$, т. е. при $n > \frac{2}{\varepsilon}$. Итак, в качестве N может быть взято любое число, не меньшее чем $\frac{2}{\varepsilon}$.

В примере 3 $a_n = \frac{1}{n^\alpha}$, и роль N может играть любое число, не меньшее чем $\varepsilon^{-\frac{1}{\alpha}}$.

В примере 4 всё обстоит так же, как в примере 1, поскольку $\left|\frac{(-1)^{n+1}}{n}\right| = \frac{1}{n}$.

В примере 5, полагая $a_n = aq^{n-1}$, мы должны иметь $|a| \cdot |q|^{n-1} < \varepsilon$, что равносильно

$$n > \frac{\lg|a| + \lg\frac{1}{\varepsilon}}{\lg\frac{1}{|q|}} + 1.$$

Таким образом, N должно быть не меньше, чем правая часть этого неравенства.

Из приведённых примеров видно, как следует понимать зависимость числа N, упоминаемого в определении предела, от числа ε. Нужно, впрочем, заметить, что на практике, если требуется найти предел данной последовательности, то прибегать к «эпсилонным» рассуждениям не приходится, так как, основываясь на теоремах, излагаемых ниже, почти всегда удаётся сводить более сложные случаи к более простым. Лишь редко подобного рода сведение оказывается невозможным, и тогда «эпсилонные» рассуждения служат средством доказательства.

Следующей теоремой придётся не раз пользоваться в дальнейшем:

Из всякой ограниченной последовательности можно «выделить» сходящуюся последовательность.

Это значит: если последовательность $\{a_n\}$ — ограниченная, то можно указать такую возрастающую последовательность целых положительных чисел $\{p_n\}$, что последовательность $\{a_{p_n}\}$ будет иметь предел.

Доказательство. По теореме Больцано-Вейерштрасса (§ 36) последовательность $\{a_n\}$ имеет хоть одну предельную точку, например a. По определению предельной точки, во всякой ε-окрестности точки a имеется точка последовательности $\{a_n\}$. Пусть $\{\varepsilon_n\}$ — последовательность положительных убывающих чисел, сходящаяся к нулю. Для каждого ε_n ($n = 1, 2, 3, \ldots$) подберём член данной последовательности a_{p_n}, находящийся в ε_n-окрестности точки a:

$$|a_{p_n} - a| < \varepsilon_n \qquad (n = 1, 2, 3, \ldots); \tag{40}$$

при этом, подбирая члены последовательности a_{p_n} один за другим, можно позаботиться, чтобы каждый следующий индекс p_{n+1} был больше предыдущего p_n. Из соотношений (40) сейчас же следует, что

$$a_{p_n} \to a.$$

Примечание. Мы рассматриваем в этом параграфе лишь ограниченные последовательности; заметим, однако, теперь же, что *неограниченная последовательность не может иметь конечного предела, т. е. если последовательность $\{a_n\}$ имеет конечный предел a, то она непременно — ограниченная.*

В самом деле, задав себе какую-нибудь ε-окрестность точки a, мы видим из соотношения (38), что все члены последовательности a_n, начиная с индекса $N+1$, удовлетворяют неравенству

$$a_n < a + \varepsilon.$$

Вместе с тем среди конечного числа членов $a_1, a_2, \ldots, a_N$ можно выбрать наибольший; обозначая затем через M любое число, превышающее как этот наибольший член, так и число $a+\varepsilon$, мы будем иметь при в с е х значениях n неравенство

$$a_n < M,$$

откуда следует ограниченность сверху. Подобным же образом устанавливается и ограниченность снизу.

На «переход к пределу» («предельный переход») можно смотреть, как на некоторую о п е р а ц и ю. С арифметическими операциями предельный переход имеет общее, но кое в чём от них заметно отличается. Как и в арифметических операциях, здесь по данным числам

$$a_1, a_2, \ldots, a_n, \ldots$$

составляется новое число $a = \lim a_n$; как и в арифметических операциях, данными числами вновь составляемое число определяется однозначно. Но операция перехода к пределу далеко не всегда «возможна» (данная последовательность чисел может не иметь предела).

Далее, «данных» чисел должно быть не два (или конечное число, как в случае сложения или умножения), а бесконечное множество; притом весьма замечательно, что в результате отбрасывания или добавления или замены к о н е ч н о г о числа членов последовательности не изменяется сходимость последовательности [1]) и не изменяется её предел (если он существует). В самом деле, если точка a есть единственная предельная точка последовательности $\{a_n\}$, то она, несомненно, будет единственной и для видоизменённой (указанным способом) последовательности.

Весьма важно, что отбрасывать можно и б е с к о н е ч н о е множество членов последовательности — лишь бы оставлено было также бесконечное множество. Именно, справедливо следующее утверждение:

Пусть

$$\{p_n\} \equiv p_1, p_2, \ldots, p_n, \ldots$$

есть некоторая возрастающая последовательность целых положительных чисел. Тогда, если предел последовательности $\{a_n\}$ *существует*

$$a_n \to a,$$

---

[1]) То есть сходящаяся последовательность остаётся сходящейся, расходящаяся — расходящейся.

то существует и равен ему предел «частной» последовательности [1] $\{a_{\rho_n}\}$:

$$a_{\rho_n} \to a.$$

В самом деле, раз последовательность $\{a_n\}$ имеет предел a (и следовательно, ограничена), то это значит, что a есть единственная предельная точка этой последовательности. Так как все члены последовательности $\{a_{\rho_n}\}$ содержатся в последовательности $\{a_n\}$, то последовательность $\{a_{\rho_n}\}$ также ограничена и вместе с тем не может иметь иных предельных точек, кроме a. Но по теореме Больцано-Вейерштрасса (§ 36) она имеет непременно хоть одну предельную точку. Итак, a есть единственная предельная точка последовательности $\{a_{\rho_n}\}$, т. е. есть её предел.

Следующие основные теоремы характеризуют свойства предельного перехода как операции и вместе с тем лежат в основе фактического вычисления пределов [2].

Теорема I. Если последовательности $\{a_n\}$ и $\{b_n\}$ имеют пределы, именно,

$$a_n \to a, \quad b_n \to b,$$

то «последовательность-сумма» $\{a_n + b_n\}$ также имеет предел, именно,

$$a_n + b_n \to a + b.$$

Теорема II. При тех же предположениях «последовательность-разность» $\{a_n - b_n\}$ имеет предел, а именно,

$$a_n - b_n \to a - b.$$

Теорема III. При тех же предположениях «последовательность-произведение» $\{a_n b_n\}$ имеет предел, а именно,

$$a_n b_n \to ab.$$

Теорема IV. При тех же предположениях и при дополнительном предположении $b \neq 0$ «последовательность-частное» $\left\{\frac{a_n}{b_n}\right\}$ имеет предел, а именно,

$$\frac{a_n}{b_n} \to \frac{a}{b}.$$

Очень часто теорему I читают кратко: «предел суммы равен сумме пределов» и записывают:

$$\lim (a_n + b_n) = \lim a_n + \lim b_n,$$

причём условия теоремы подразумеваются. Если пределы a_n и b_n не существуют, то, конечно, в общем случае («раз навсегда») нельзя ничего сказать

---

[1]) Вместо «частная последовательность» говорят также «подпоследовательность».

[2]) Доказательства этих теорем были приведены на стр. 194–195 кн. 1 «Энциклопедии элементарной математики». Мы приводим, однако, несколько иные приёмы доказательств.

о пределе a_n+b_n. Следует, впрочем, заметить, что возможны случаи, когда оба предела a_n и b_n не существуют, а предел a_n+b_n существует [1]).

Аналогичные замечания справедливы и по поводу теорем II — IV.

Докажем сначала теорему I. Нам дано, что при достаточно больших значениях n отклонение $|a_n-a|$ делается сколь угодно малым, и то же относительно $|b_n-b|$; требуется доказать то же самое по поводу отклонения $|(a_n+b_n)-(a+b)|$.

Мы имеем по свойству абсолютной величины [2])

$$|(a_n+b_n)-(a+b)|=|(a_n-a)+(b_n-b)|\leqslant|a_n-a|+|b_n-b|.$$

Пусть задано наперёд число ε (>0). Тогда можно подобрать такое N'_ε, что при $n>N'_\varepsilon$

$$|a_n-a|<\frac{\varepsilon}{2},$$

и такое N''_ε, что при $n>N''_\varepsilon$

$$|b_n-b|<\frac{\varepsilon}{2}.$$

В таком случае, обозначая через N_ε наибольшее из чисел N'_ε и N''_ε, будем иметь при $n>N_\varepsilon$:

$$|(a_n+b_n)-(a+b)|\leqslant|a_n-a|+|b_n-b|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

что и требовалось доказать.

Докажем теперь теорему III. При тех же данных, что и в теореме I, требуется доказать, что при достаточно больших значениях n отклонение $|a_nb_n-ab|$ станет меньше любого наперёд заданного ε (>0). Мы имеем:

$$a_nb_n-ab=(a_n-a)\,b+a\,(b_n-b)+(a_n-a)\,(b_n-b),$$

и по свойствам абсолютной величины

$$|a_nb_n-ab|\leqslant$$
$$\leqslant|a_n-a|\cdot|b|+|a|\cdot|b_n-b|+|a_n-a|\cdot|b_n-b|. \quad (41)$$

---

[1]) Например, последовательности

$$1,\ -1,\ 1,\ -1,\ 1,\ -1,\ \ldots$$

и

$$-1,\ 1,\ -1,\ 1,\ -1,\ 1,\ \ldots$$

расходящиеся, тогда как «последовательность-сумма»

$$0,\ 0,\ 0,\ 0,\ 0,\ 0,\ \ldots$$

сходящаяся.

[2]) Мы имеем в виду свойство

$$|x+y|\leqslant|x|+|y|.$$

Далее используется ещё свойство

$$|xy|=|x|\cdot|y|.$$

См. Э.э.м., кн. 1, формулы (1) и (2) на стр. 128—129.

Пусть K — наибольшее из чисел $|a|$, $|b|$ и 1[1]); достаточно, конечно, допустить по поводу ε, что

$$\varepsilon < 3K^2.$$

Тогда, подобрав N'_ε и N''_ε таким образом, что

$$|a_n - a| < \frac{\varepsilon}{3K} \quad \text{при} \quad n > N'_\varepsilon$$

и

$$|b_n - b| < \frac{\varepsilon}{3K} \quad \text{при} \quad n > N''_\varepsilon,$$

мы будем иметь при $n > N_\varepsilon$, где $N_\varepsilon = \max\{N'_\varepsilon, N''_\varepsilon\}$:

$$|a_n - a| \cdot |b| + |a| \cdot |b_n - b| + |a_n - a| \cdot |b_n - b| \leqslant$$

$$\leqslant \frac{\varepsilon}{3K} \cdot K + K \cdot \frac{\varepsilon}{3K} + \frac{\varepsilon}{3K} \cdot \frac{\varepsilon}{3K} = \varepsilon\left(\frac{2}{3} + \frac{\varepsilon}{9K^2}\right) < \varepsilon\left(\frac{2}{3} + \frac{3K^2}{9K^2}\right) = \varepsilon. \quad (42)$$

Сопоставляя (41) и (42), получаем то, что требовалось доказать.

Отметим частный случай теоремы III (при $\{a_n\} \equiv \{a\}$): *если* $b_n \to b$, *то* $ab_n \to ab$ («постоянный множитель можно выносить за знак предела»). В частности, при $a = -1$ получается: *если* $b_n \to b$, *то* $(-b_n) \to (-b)$. Отсюда сейчас же вытекает теорема II

$$a_n - b_n = a_n + (-b_n) \to a + (-b) = a - b.$$

Переходя к теореме IV[2]), докажем сначала её частный случай, когда $\{a_n\} \equiv \{1\}$: если $b_n \to b \neq 0$, то $\frac{1}{b_n} \to \frac{1}{b}$.

Составим отклонение $\frac{1}{b_n}$ от $\frac{1}{b}$:

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b_n - b|}{|b_n| \cdot |b|}. \quad (43)$$

Отклонение $|b_n - b|$ при достаточно больших значениях n становится меньше любого положительного числа; например, оно станет меньше, чем $\frac{|b|}{2}$:

$$|b_n - b| < \frac{|b|}{2} \qquad (\text{при } n > N').$$

Но

$$|b| = |(b - b_n) + b_n| \leqslant |b - b_n| + |b_n|,$$

так что

$$|b - b_n| \geqslant |b| - |b_n|.$$

---

[1]) Последнее нужно на случай, если $a = b = 0$.

[2]) Заметим, между прочим, что текст теоремы IV нуждается в дополнительном разъяснении: из неравенства $b \neq 0$ вытекает, что при достаточно больших n величины b_n отличны от нуля; таким образом, последовательность $\left\{\frac{1}{b_n}\right\}$ может содержать лишь конечное число членов, не имеющих смысла; их следует отбросить или заменить какими угодно числами.

Отсюда следует:

$$|b| - |b_n| < \frac{|b|}{2},$$

т. е.

$$|b_n| > \frac{1}{2}|b|\ ^{1}). \tag{44}$$

Подберём N_ε'' так, чтобы из неравенства $n > N_\varepsilon''$ следовало:

$$|b_n - b| < \frac{1}{2}|b|^2\varepsilon. \tag{45}$$

Тогда из сопоставления (43), (44) и (45) вытекает при $n > N_\varepsilon$, где $N_\varepsilon = \max\{N', N_\varepsilon''\}$:

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| < \frac{\frac{1}{2}|b|^2 \cdot \varepsilon}{|b| \cdot \frac{1}{2}|b|} = \varepsilon,$$

что и требовалось доказать.

Теперь общий случай теоремы IV сводится к теореме III: если $a_n \to a$, $b_n \to b$, то

$$\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n} \to a \cdot \frac{1}{b} = \frac{a}{b}.$$

Докажем, наконец, теорему о предельном переходе в неравенстве:

Если последовательности $\{a_n\}$ *и* $\{b_n\}$ *имеют пределы*

$$a_n \to a, \quad b_n \to b,$$

и, кроме того, при всех значениях n *справедливо неравенство*

$$a_n \leqslant b_n, \tag{46}$$

то непременно имеет место неравенство

$$a \leqslant b. \tag{47}$$

Начнём с частного случая, когда $\{a_n\} \equiv \{0\}$, и значит, $a = 0$. Нужно доказать, что из соотношений $b_n \geqslant 0$ $(n = 1, 2, 3, \ldots)$ и $b_n \to b$ следует $b \geqslant 0$. Будем доказывать от противного. Пусть $b < 0$. Полагая $\varepsilon = |b|$, видим, что по определению предела при достаточно больших n должно иметь место неравенство

$$-|b| < |b_n - b| < |b|,$$

или

$$b - |b| < b_n < b + |b|.$$

Но так как $b < 0$, то $b + |b| = 0$, и вторая половина последнего неравенства принимает вид $b_n < 0$. Это противоречит допущению $b_n \geqslant 0$ $(n = 1, 2, \ldots)$.

---

[1]) Не представит труда истолковать предыдущее рассуждение геометрически, с помощью числовой оси.

Перейдём к общему случаю. По предположению $b_n \geqslant a_n$ ($n = 1, 2, \ldots$); значит, $b_n - a_n \geqslant 0$ ($n = 1, 2, \ldots$). На основании теоремы II $b_n - a_n \to b - a$. По уже доказанному (частный случай) $b - a \geqslant 0$, т. е. $b \geqslant a$.

В частности, заключение теоремы сохраняется и в том, особенно часто встречающемся случае, когда условие (46) заменяется следующим:

$$a_n < b_n.$$

Иными словами, из неравенства $a_n < b_n$ вытекает, что $a \leqslant b$ (но не $a < b$).

Этими соображениями оправдывается следующее правило: «при переходе к пределу в неравенстве нужно добавлять знак равенства».

§ 39. Обобщение понятия предела (пределы в «несобственном смысле»)

Мы познакомимся теперь с некоторыми условными оборотами речи, широко употребляемыми в математической практике.

Условимся называть «окрестностью точки $+\infty$» всякий промежуток, не ограниченный сверху (без начальной точки), и будем такой промежуток

$$x > M \quad (\text{или } M < x < +\infty)$$

называть M-окрестностью точки $+\infty$.

Говорят, что последовательность имеет «предельную точку $+\infty$», если во всякой «окрестности точки $+\infty$» содержится бесконечное множество точек последовательности. Это как раз означает, что последовательность неограничена сверху.

Аналогично определяется «окрестность точки $-\infty$» — как всякий промежуток, не ограниченный снизу (без конечной точки); такой промежуток

$$x < -M \quad (\text{или } -\infty < x < -M)$$

называется «($-M$)-окрестностью точки $-\infty$». Если говорят, что последовательность «имеет предельную точку $-\infty$», то это равносильно констатации того, что последовательность неограничена снизу.

Формулировка теоремы Больцано-Вейерштрасса теперь упрощается:

Всякая бесконечная последовательность $\{a_n\}$ *имеет хотя бы одну предельную точку* (в «собственном» или в «несобственном» смысле).

В самом деле, если последовательность не ограничена сверху, то она имеет предельную точку $+\infty$; если не ограничена снизу, то имеет предельную точку $-\infty$; если ограничена сверху и снизу, то

по основной теореме имеет хоть одну «конечную» предельную точку.

Далее, остаётся в силе (расширяя свой смысл) определение: *пределом последовательности называется предельная точка, если она — единственная.* Допустим, что единственная предельная точка есть $+\infty$, и перефразируем это утверждение. Так как $-\infty$ не есть предельная точка, то последовательность ограничена снизу, так что в некотором промежутке $(-\infty, m)$ нет ни одной точки последовательности; так как никакая «конечная» точка не есть предельная точка, то во всяком «конечном» промежутке $[m, M]$, где $M > m$, имеется лишь конечное число точек последовательности. Среди индексов этих точек возьмём наибольший и обозначим его через $N \equiv N_M$; тогда окажется, что при условии $n > N$ точка a_n непременно находится в M-окрестности точки $+\infty$, т. е. удовлетворяет неравенству $a_n > M$. Ясно и обратное: если, как бы велико ни было M, все точки a_n при достаточно больших значениях n находятся в M-окрестности точки $+\infty$, то последовательность $\{a_n\}$ не имеет иных предельных точек, кроме $+\infty$. Мы приходим, таким образом, к следующему определению:

Считается, что *последовательность* $\{a_n\}$ *имеет предел* $+\infty$ (*«a_n стремится к бесконечности»*), *если, как бы велико ни было наперёд заданное число* M, *можно указать такое число* $N \equiv N_M$, *что неравенство*

$$n > N \tag{48}$$

влечёт за собой неравенство

$$a_n > M. \tag{49}$$

Иначе: *последовательность* $\{a_n\}$ *имеет предел* $+\infty$, *если её общий член становится сколь угодно большим, раз только достаточно велик его индекс.*

Пишут:

1) $\lim a_n = \infty$,

2) $a_n \to \infty$.

Иногда вместо того, чтобы говорить «стремится к бесконечности», говорят «расходится к бесконечности».

Воздерживаясь от повторений, скажем, что последовательность $\{a_n\}$ имеет предел $-\infty$, если, как бы велико ни было наперёд заданное число M, можно указать такое число $N = N_M$, что неравенство

$$n > N \tag{50}$$

влечёт за собой неравенство

$$a_n < -M. \tag{51}$$

Запись: $\lim a_n = -\infty$, или $a_n \to -\infty$.

Приведём несколько примеров.

1. Последовательность натуральных чисел $\{n\}$ имеет единственную предельную точку, т. е. предел, $+\infty$.

То же справедливо относительно последовательностей $\{n^\alpha\}$ при $\alpha > 0$.

2. Последовательность

$$\{(-1)^{n+1}n\} = 1, -2, 3, -4, 5, -6, \ldots$$

имеет две предельные точки $+\infty$ и $-\infty$, и следовательно, не имеет предела даже и в «несобственном» смысле.

3. Если последовательность $\{a_n\}$ не имеет конечных предельных точек, то $\lim |a_n| = +\infty$.

4. Арифметическая прогрессия $\{a + (n-1)d\}$ имеет предел $+\infty$ или $-\infty$, смотря по тому, будет ли $d > 0$, или $d < 0$.

5. Геометрическая прогрессия $\{q^n\}$ имеет предел $+\infty$ при $q > 1$ [1]), и не имеет предела при $q < -1$. Но в этом последнем случае, конечно, $\lim |q|^n = +\infty$.

6. $\lim \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right) = \infty$ (см. конец § 35).

7. Последовательность 2 на стр. 151 имеет в качестве предельных точек все натуральные числа и ещё $+\infty$. Предела нет.

8. Последовательность $\{\operatorname{tg} n\pi\theta\}$, где θ — иррациональное число, имеет предельными точками все числа без исключения и ещё $+\infty$ и $-\infty$.

В самом деле, если бы в некотором промежутке (α, β) не было ни одной точки последовательности $\{\operatorname{tg} n\pi\theta\}$, то в промежутке $\left(\frac{\alpha}{\sqrt{1+\alpha^2}}, \frac{\beta}{\sqrt{1+\beta^2}}\right)$ не было бы ни одной точки последовательности $\{\sin n\pi\theta\}$, а это противоречит примеру 16 § 37.

По поводу основных теорем I—IV § 38 необходимо сделать предостережение: *на случай бесконечных пределов они переносятся лишь частично.* Справедливы такие теоремы (доказательство их предоставляем читателю):

I'. Если $a_n \to +\infty$, $b_n > m$, то $a_n + b_n \to +\infty$.

I''. Если $a_n \to -\infty$, $b_n < M$, то $a_n + b_n \to -\infty$.

(Теоремы типа II заменой знака сводятся к теоремам типа I).

III'. Если $a_n \to \infty$, $b_n > m > 0$, то $a_n b_n \to \infty$.

III''. Если $a_n \to 0$, $|b_n| < M$, то $a_n b_n \to 0$.

IV'. Если $a_n \to \infty$, $0 < b_n < M$, то $\frac{a_n}{b_n} \to \infty$.

IV''. Если $a_n \to 0$, $|b_n| > m > 0$, то $\frac{a_n}{b_n} \to 0$.

IV'''. Если $|a_n| < M$, $|b_n| \to \infty$, то $\frac{a_n}{b_n} \to 0$.

IV''''. Если $|a_n| > m > 0$, $|b_n| \to 0$, то $\left|\frac{a_n}{b_n}\right| \to \infty$.

[1]) Доказательство см. на стр. 81.

Но нельзя ответить в общей форме («раз навсегда»), например на такие вопросы:

1) Что будет с разностью $a_n - b_n$, если $a_n \to \infty$, $b_n \to \infty$?

2) Что будет с произведением $a_n b_n$, если $a_n \to \infty$, $b_n \to 0$?

3) Что будет с отношением $\frac{a_n}{b_n}$, если $a_n \to \infty$, $b_n \to \infty$?

4) Что будет с отношением $\frac{a_n}{b_n}$, если $a_n \to 0$, $b_n \to 0$?

Каждый из вопросов подобного рода получает разрешение только в том случае, если указано, каковы именно данные последовательности $\{a_n\}$ и $\{b_n\}$. Очень часто простое тождественное преобразование приводит к положениям, когда одна из теорем типов I—IV уже оказывается применимой [1]); но иногда всё же приходится прибегать к особому исследованию (примеры см. ниже, §§ 43—45).

§ 40. Предел функции на бесконечности

К понятию предела последовательности сводится понятие предела функции. Рассмотрим сначала случай предела функции «на бесконечности» ($+\infty$).

Предположим, что некоторая функция $f(x)$ действительной переменной x задана (не теряет смысла) в некоторой окрестности точки ($+\infty$), например в промежутке (a, ∞). Мы уже имели дело с последовательностями, общий член которых a_n имеет вид $f(x_n)$, где $x_n = n$:

$$\{f(n)\} \equiv f(1), f(2), \ldots, f(n), \ldots$$

Пусть теперь $\{x_n\}$ — какая угодно последовательность значений x из промежутка (a, ∞), имеющая предел $+\infty$:

$$x_n \to +\infty. \tag{52}$$

Если последовательность соответствующих значений функции

$$\{f(x_n)\} \equiv f(x_1), f(x_2), \ldots, f(x_n), \ldots$$

также стремится к некоторому пределу L, конечному или бесконечному [2]),

$$f(x_n) \to L, \tag{53}$$

и притом независимо от того, как выбрана последовательность $\{x_n\}$, обладающая свойством (52), то тогда говорят, что *при не-*

---

[1]) Распространённый термин «раскрытие неопределённостей» (случаи: $\infty - \infty$, $\infty \cdot 0$, $\frac{\infty}{\infty}$, $\frac{0}{0}$) способен вносить некоторую дезориентацию, так как «неопределённость» заключается лишь в невозможности непосредственного применения какой-либо из основных теорем.

[2]) Условимся пользоваться большой буквой L для того, чтобы охватить сразу оба эти случая.

ограниченном возрастании x функция стремится к пределу L, и записывают это одним из следующих способов:

$$\left.\begin{array}{l} 1)\ \lim\limits_{x\to\infty} f(x) = L, \\ 2)\ f(x)\to L \text{ при } x\to+\infty. \end{array}\right\} \qquad (54)$$

Напротив, если две различные последовательности вида $\{f(x_n)\}$ имеют различные пределы или существует хотя бы одна такая последовательность $\{x_n\}$, что последовательность $\{f(x_n)\}$ не имеет предела [1]), то говорят, что функция $f(x)$ *не стремится к пределу* при неограниченном возрастании x, или что предел $\lim\limits_{x\to+\infty} f(x)$ *не существует.*

Подобные же определения даются понятию предела $f(x)$ при $x\to-\infty$.

Рассмотрим примеры.

1. Функция $f(x)=\frac{1}{x}$ при $x\to\infty$ имеет предел 0:

$$\lim_{x\to\infty}\frac{1}{x}=0.$$

В самом деле, раз $x_n\to\infty$, то $\frac{1}{x_n}\to 0$ (IV″). Точно так же

$$\lim_{x\to-\infty}\frac{1}{x}=0.$$

2. Пусть $f(x)=x^n$ (n — целое положительное). Тогда

$$\lim_{n\to\infty} x^n=+\infty,\quad \lim_{x\to-\infty} x^n=\begin{cases}+\infty, & \text{если } n \text{ — чётное,}\\ -\infty, & \text{если } n \text{ — нечётное.}\end{cases} \qquad (III')$$

3. Пусть $f(x)=a^x$ ($a>1$). Тогда

$$\lim_{x\to\infty} a^x=\infty,\quad \lim_{x\to-\infty} a^x=0.$$

Для целых значений x это показано на стр. 81; для произвольных — следует из того, что функция a^x — возрастающая.

4. $\lim\limits_{x\to\infty}\log_a x=\infty$ (при $a>1$).

5. Функция $f(x)=\sin x$ (рис. 70, a) не имеет предела при x, стремящемся к ∞. Действительно, положив $x_n=\pi n$, мы получаем $f(\pi n)=\sin\pi n\to 0$; положив $x_n=\frac{\pi}{2}+2\pi n$, получаем

---

[1]) Случай, когда всевозможные последовательности вида $\{f(x_n)\}$ имеют предел, но не всегда один и тот же, является невозможным. Именно, если последовательность $\{f(x'_n)\}$ имеет предел L', а последовательность $\{f(x''_n)\}$ — предел L'', причём $L'\neq L''$, то последовательность

$$f(x'_1),\ f(x''_2),\ f(x'_3),\ f(x''_4),\ \ldots,\ f(x'_{2n-1}),\ f(x''_{2n}),\ldots$$

предела иметь не может (см. § 37, пример 9).

$f\left(\frac{\pi}{2}+2\pi n\right)=\sin\left(\frac{\pi}{2}+2\pi n\right)\to 1$; положив $x_n=\frac{3}{2}\pi+2\pi n$, получаем $f\left(\frac{3}{2}\pi+2\pi n\right)=\sin\left(\frac{3}{2}\pi+2\pi n\right)\to -1$, и т. п.

Аналогично не существует предел $\lim\limits_{x\to-\infty}\sin x$.

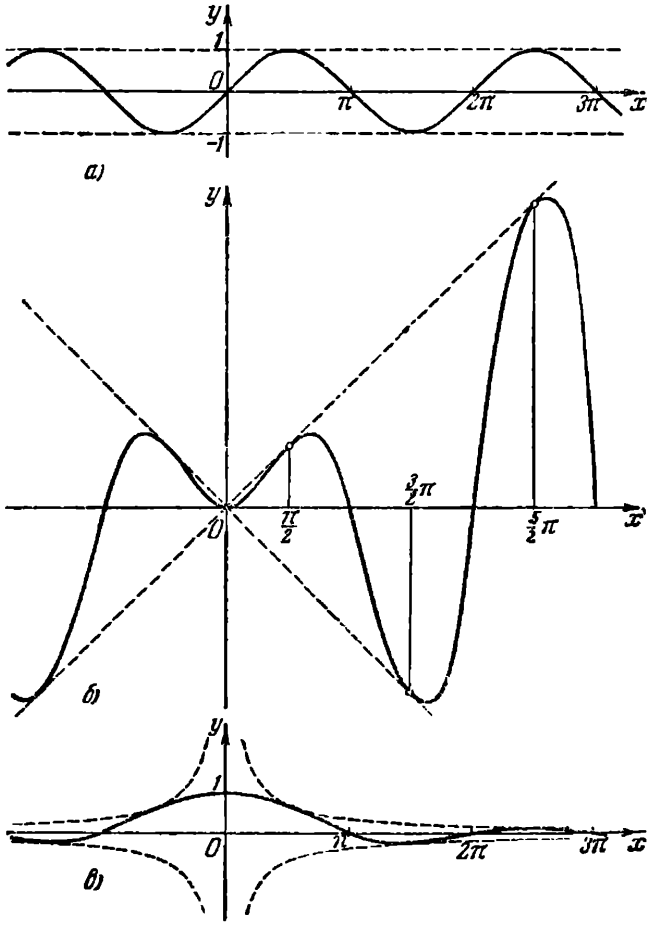


Рис. 70.

6. Функция $f(x)=x\sin x$ (рис. 70, *б*) также не имеет предела ни при $x\to+\infty$, ни при $x\to-\infty$. На этот раз, как в предыдущем примере, $f(\pi n)\to 0$; но $f\left(\frac{\pi}{2}+2\pi n\right)\to+\infty$, $f\left(\frac{3}{2}\pi+2\pi n\right)\to-\infty$.

7. Функция $f(x)=\frac{\sin x}{x}$ (рис. 70, *в*) имеет предел 0 при $x\to+\infty$ и при $x\to-\infty$:

$$\lim_{x\to\pm\infty}\frac{\sin x}{x}=0.$$

Это ясно из неравенства

$$\left|\frac{\sin x}{x}\right| = \frac{|\sin x|}{|x|} \leqslant \frac{1}{|x|},$$

с использованием примера 1.

8. Функция $f(x) = 2^x \sin 2\pi x$ не имеет предела при $x \to +\infty$ и имеет предел 0 при $x \to -\infty$:

$$\lim_{x \to -\infty} 2^x \sin 2\pi x = 0.$$

Функция $f(x) = 2^{-x} \sin 2\pi x$ (рис. 71), напротив, имеет предел 0 при $x \to +\infty$:

$$\lim_{x \to +\infty} 2^{-x} \sin 2\pi x = 0,$$

и не имеет предела при $x \to -\infty$.

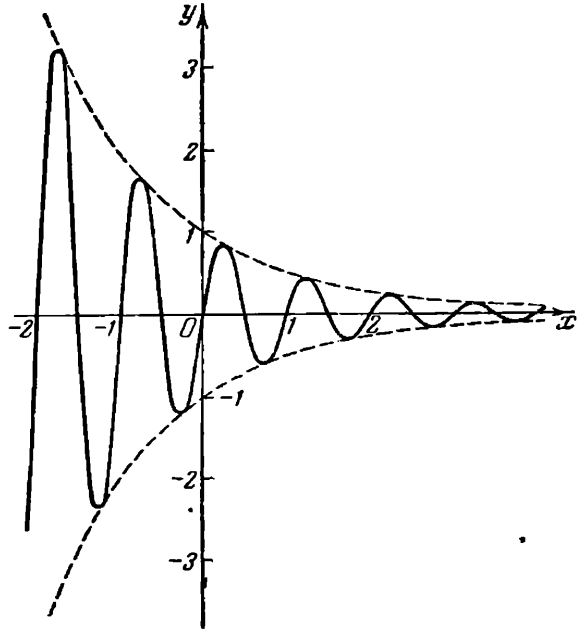


Рис. 71.

Определению «предела функции на бесконечности» можно дать иную («эпсилонную») формулировку. Остановимся на случае точки $+\infty$. Соотношение (54) означает:

1) При L конечном:

как бы мало ни было $\varepsilon\,(>0)$, *можно указать такое число* $X \equiv X_\varepsilon$, *что из неравенства*

$$x > X$$

следует неравенство

$$|f(x) - L| < \varepsilon; \quad (55)$$

2) при $L = +\infty$:

как бы велико ни было M, *можно указать такое число* $X \equiv X_M$, *что из неравенства*

$$x > X$$

следует неравенство

$$f(x) > M; \quad (56)$$

3) при $L = -\infty$:

такая же формулировка, с заменой неравенства (56) неравенством

$$f(x) < -M. \quad (57)$$

Покажем, что новые и прежние формулировки равносильны, останавливаясь хотя бы на случае L конечного.

Допустим, что из $x_n \to \infty$ следует $f(x_n) \to L$, и докажем, что ко всякому $\varepsilon\,(>0)$ можно подобрать число $X \equiv X_\varepsilon$, обладающее

указанным свойством. Пусть это неверно, и такое число X можно подобрать не ко всякому ε; предположим, что к некоторому $\varepsilon^*\,(>0)$ нельзя подобрать числа X, обладающего требуемым свойством. Это значит, что, как бы велико ни было X, можно к нему подобрать такое значение x, что $x>X$, но $|f(x)-L|\geqslant\varepsilon^*$. Пусть дана последовательность

$$\{X_n\}\equiv X_1,\ X_2,\ \ldots,\ X_n,\ \ldots$$

такая, что $X_n\to\infty$. К каждому X_n подберём такое $x\equiv x_n$, что

$$x_n>X_n, \tag{58}$$

но вместе с тем

$$|f(x_n)-L|\geqslant\varepsilon^*. \tag{59}$$

Рассмотрим теперь последовательность значений x

$$\{x_n\}\equiv x_1,\ x_2,\ \ldots,\ x_n,\ \ldots$$

Из неравенства (58) следует, что $x_n\to\infty$. Но из неравенства (59) следует, что соотношение $f(x_n)\to L$ не осуществляется, так как отклонение $|f(x_n)-L|$ остаётся большим, чем положительное число ε^*. Получается противоречие, откуда ясно, что сделанное допущение было неверно.

Обратно, допустим, что ко всякому $\varepsilon\,(>0)$ можно подобрать число $X\equiv X_\varepsilon$, обладающее тем свойством, что из неравенства $x>X$ следует неравенство (55), и покажем, что из $x_n\to\infty$ следует $f(x_n)\to L$. Так как $x_n\to\infty$, то, как бы велико ни было M, при $n>N$, где $N\equiv N_M$, мы должны иметь: $x_n>M$. Положим $M=X$; тогда при $n>N$ будем иметь $x_n>X$, и в таком случае согласно сделанному допущению отсюда вытекает $|f(x_n)-L|<\varepsilon$.

§ 41. Односторонний предел функции в конечной точке

Перейдём к рассмотрению вопроса о пределе значений функции $f(x)$ в некоторой конечной точке $x=c$, т. е. о пределе, к которому, возможно, стремится значение $f(x)$ при неограниченном приближении переменной x к конечному значению c.

Начнём с примеров, демонстрирующих разнообразие возникающих в данном случае возможностей.

Пример 1. Исследуем поведение функции

$$f(x)=2^{\frac{1}{x}}, \tag{60}$$

в особенности сосредоточивая внимание на окрестности точки $x=0$. Функция $f(x)$ задаётся предыдущей формулой для всех значений x, кроме $x=0$; очевидно, при всех значениях $x\,(\neq 0)$ она положительна. Легко понять, что

$$\lim_{x\to+\infty}f(x)=\lim_{x\to-\infty}f(x)=1.$$

Если x убывает от $+\infty$, неограниченно приближаясь к нулю, то $\frac{1}{x}$ возрастает от 0 до бесконечности, и тогда $2^{\frac{1}{x}}$ возрастает от 1 до бесконечности. С другой стороны, если x возрастает от $-\infty$, приближаясь к 0, то $\frac{1}{x}$ убывает от 0 до $-\infty$, и значит, $2^{\frac{1}{x}}$ убывает от 1 до 0 (причём значение 0 не принимается ни при каком значении x).

График функции (60) показан на рис. 72.

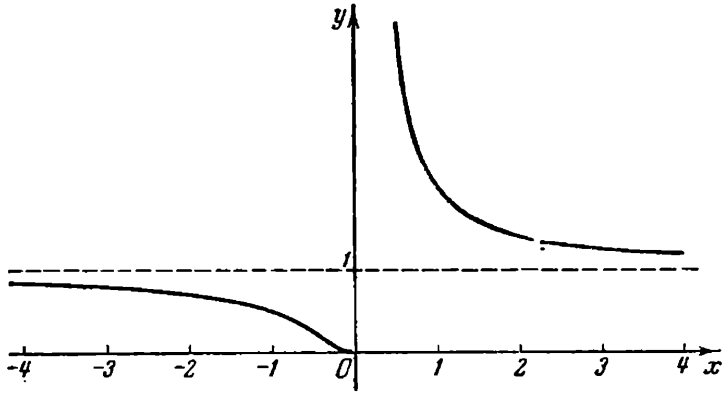


Рис. 72.

Замечательное свойство функции (60), которое нас интересует, заключается в наличии разрыва («скачка») её графика около точки $x=0$; чуть правее этой точки значения функции чрезвычайно велики, чуть левее — близки к нулю, в самой же точке $x=0$ функция не имеет никакого значения: на вертикальной оси нет ни одной точки графика.

Пример 2.

$$f(x)=\frac{1}{1+2^{\frac{1}{x}}}. \qquad (61)$$

Этот более сложный пример исследуется с точки зрения возрастания и убывания подобно предыдущему и приводит к графику, изображённому на

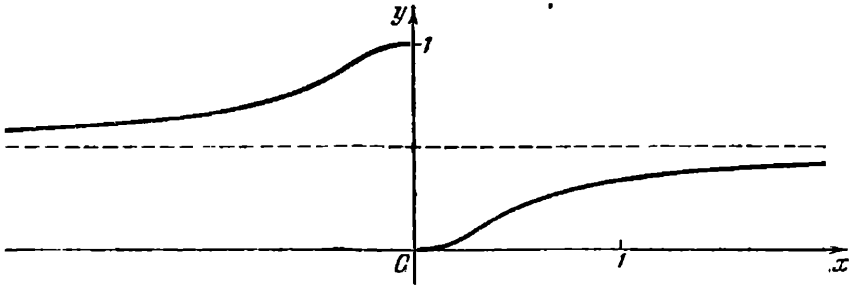


Рис. 73.

рис. 73. Здесь также имеется разрыв в точке $x=0$; но тогда как чуть левее этой точки значения $f(x)$ близки к 1, чуть правее этой точки они близки

к нулю («конечный скачок»), и опять на вертикальной оси нет ни одной точки графика, так как функция не определена при $x=0$.

Функция, между прочим, обладает свойством

$$f(x)+f(-x)=1,$$

откуда можно заключить, что точка $\left(0, \frac{1}{2}\right)$ есть центр симметрии графика.

Пример 3. Достаточно в данной связи лишь напомнить читателю о функциях

$$\text{а) } f(x)=\frac{1}{x} \quad \text{и} \quad \text{б) } f(x)=\frac{1}{x^2},$$

графики которых обнаруживают разрывы при приближении к точке $x=0$: с обеих сторон имеется неограниченное возрастание по абсолютной величине, но в первом случае — при различных знаках, во втором — при одном и том же знаке.

Пример 4.

$$f(x)=\sin\frac{1}{x}. \tag{62}$$

Значения функции, очевидно, колеблются между $+1$ и -1. Ясно, что $\lim\limits_{x\to\infty} f(x)=0$.

Когда x убывает от $+\infty$ до $\frac{2}{\pi}$, то $\frac{1}{x}$ возрастает от 0 до $\frac{\pi}{2}$ и $f(x)$ возрастает от 0 до 1; когда же x убывает от $\frac{2}{\pi}$ до $\frac{2}{3\pi}$, то $\frac{1}{x}$ возрастает от $\frac{\pi}{2}$ до $\frac{3}{2}\pi$ и $f(x)$ убывает от 1 до -1, обращаясь в нуль при $x=\frac{1}{\pi}$, и т. д.

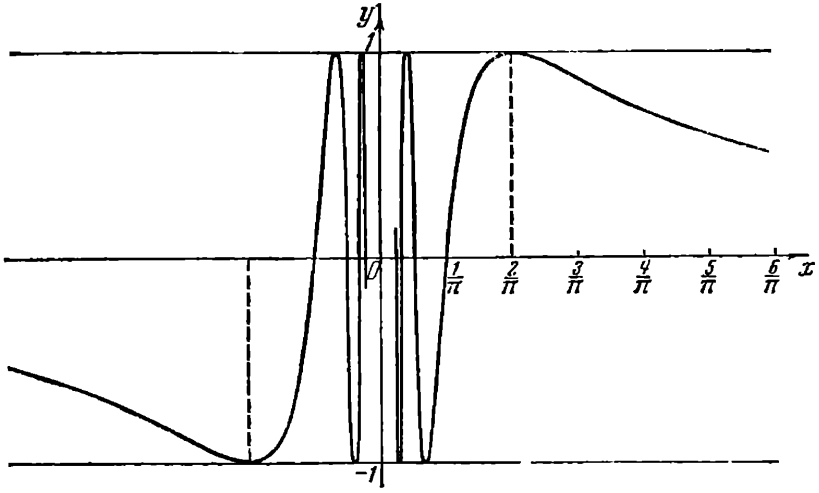


Рис. 74.

Так как функция $f(x)$ — нечётная, то её график, изображённый на рис. 74, имеет начало O центром симметрии. При приближении x к нулю колебания функции «учащаются». В точке $x=0$ функция не определена.

Пример 5.

$$f(x)=\frac{1}{x}\sin\frac{1}{x}. \tag{63}$$

Функция «совершает колебания», которые «учащаются» и «растут» (в смысле размаха) при приближении x к точке O. Так как при $x>0$

$$|f(x)|\leqslant\frac{1}{x},$$

т. е.

$$-\frac{1}{x}\leqslant f(x)\leqslant+\frac{1}{x},$$

то график заключён между ветвями гипербол $y=\frac{1}{x}$ и $y=-\frac{1}{x}$ (рис. 75). Общие точки графика с верхней гиперболой получаются из уравнения

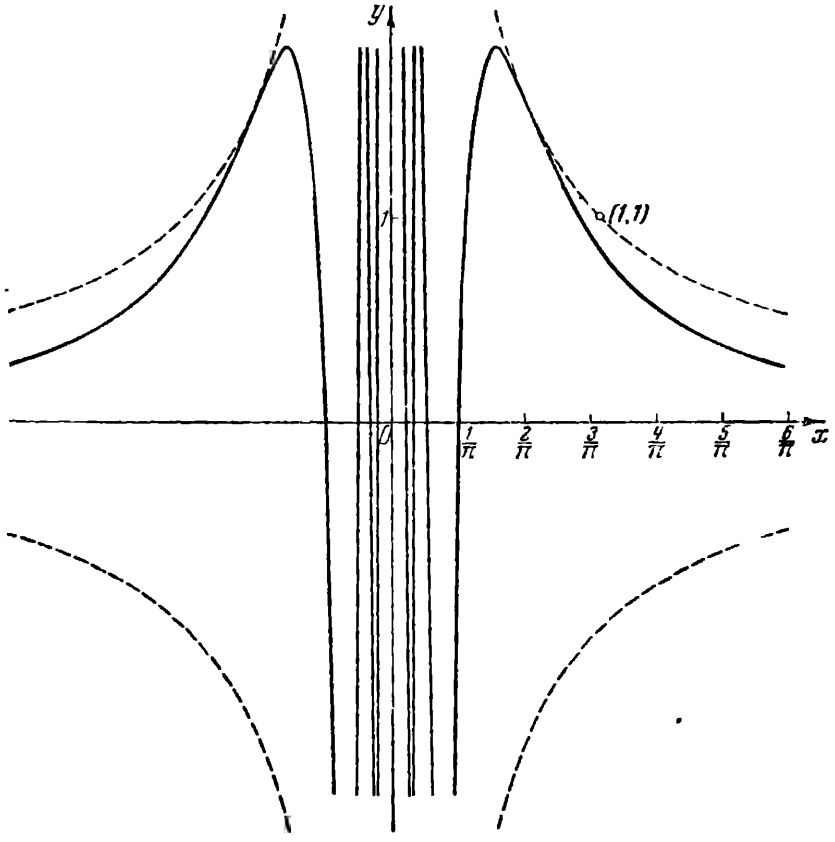


Рис. 75.

$\sin\frac{1}{x}=1$, что даёт $x=\frac{2}{(2k+1)\pi}$; общие точки с нижней гиперболой соответствуют значениям $x=\frac{2}{(2k+3)\pi}$ $\left(\text{корни уравнения } \sin\frac{1}{x}=-1\right)$. Функция — чётная, график симметричен относительно оси Oy. С осью Oy график не пересекается; функция не имеет смысла при $x=0$.

Пример 6.

$$f(x)=x\sin\frac{1}{x}. \tag{64}$$

В этом примере функция также совершает колебания, которые на этот раз отличаются тем, что их «размах» уменьшается при приближении к началу O. При $x>0$, очевидно,

$$|f(x)|\leqslant x,$$

так что

$$-x\leqslant f(x)\leqslant +x. \tag{65}$$

Равенство справа достигается при $x=\frac{2}{(2k+1)\pi}$, равенство слева — при $x=\frac{2}{(2k+3)\pi}$. Максимумы и минимумы находятся чуть правее этих точек [1].

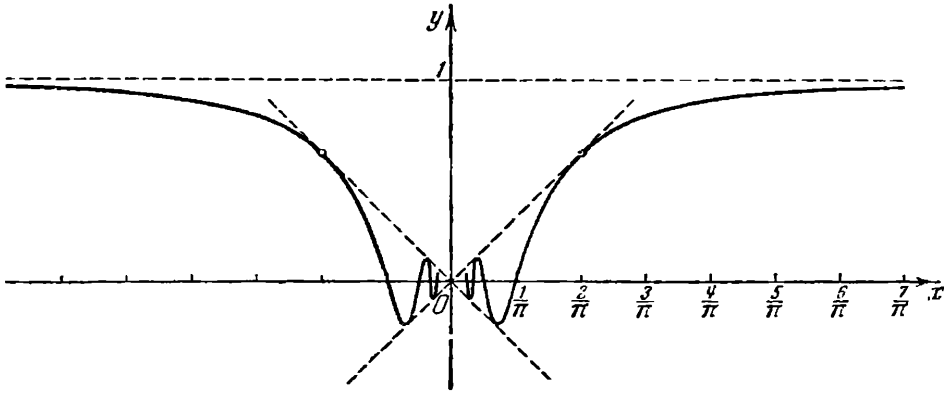


Рис. 76.

Функция — чётная, график имеет ось симметрии Oy (рис. 76). Несмотря на неравенство (65), разрыв не устранён окончательно; в точке $x=0$ функция не имеет смысла, и на оси Oy нет точек графика [2].

---

[1]) Легче всего это установить с помощью дифференциального исчисления. Мы получим:

$$f'(x)=\sin\frac{1}{x}-\frac{1}{x}\cos\frac{1}{x}.$$

Отсюда следует:

$$f'\left(\frac{2}{(2k+1)\pi}\right)=(-1)^k\neq 0.$$

Решить же уравнение $f'(x)=0$ можно, приведя его к виду $\operatorname{tg}\frac{1}{x}=\frac{1}{x}$, или $\operatorname{tg}\xi=\xi$, где $\xi=\frac{1}{x}$, а затем действуя графически.

[2]) Положим $\frac{1}{x}=u$. Так как $\sin u<u$ при $u>0$, то $x\sin\frac{1}{x}<1$ при $x>0$. Далее, из соотношения

$$\lim_{u\to 0}\frac{\sin u}{u}=1$$

Графики примеров 5 и 6 удобно строить систематическими приёмами, указанными в § 2.

Пример 7.

$$f(x)=2^{-\frac{1}{x^2}}. \tag{66}$$

Функция не определена при $x=0$; однако при приближении к значению $x=0$ справа или слева без «колебаний» и очень быстро приближается к нулю (рис. 77).

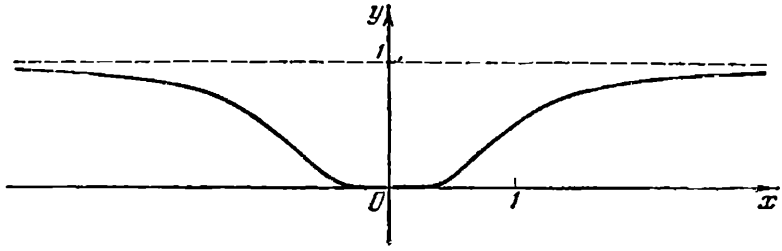


Рис. 77.

Примеры, подобные рассмотренным, дают повод ввести следующие уточняющие определения.

Пусть функция $f(x)$ задана в некотором промежутке с конечной точкой c, кроме самой точки c, в которой она может быть задана или не задана. Если для всякой последовательности точек $\{x_n\}$, обладающих свойствами

$$\left.\begin{array}{l}1)\ x_n<c\\ 2)\ x_n\to c,\end{array}\right\}\ (n=1,\ 2,\ \ldots), \tag{67}$$

последовательность $\{f(x_n)\}$ стремится к одному и тому же конечному или бесконечному пределу L',

$$f(x_n)\to L', \tag{68}$$

то говорят, что *при неограниченном приближении x к точке c слева функция $f(x)$ стремится к пределу L'*; для записи этого пользуются иногда условным обозначением [1])

$$f(c-0)=L'.$$

(см. формулу (89) на стр. 192) следует:

$$\lim_{x\to\infty} x\sin\frac{1}{x}=1.$$

Таким образом, прямая $y=1$ есть асимптота графика рассматриваемой функции (см. пунктирную прямую на рис. 76); сам график приближается к ней снизу.

[1]) Принимается во внимание громоздкость записей вроде

$$\text{«}\lim_{\substack{x<c\\ x\to c}} f(x)=L'\text{»},\quad \text{или}\quad \text{«}f(x)\to L'\ \text{при}\ x<c,\ x\to c\text{»}.$$

Точно так же, если для всякой последовательности точек $\{x_n\}$, взятых в некотором промежутке с начальной точкой c и обладающих свойствами

$$\left.\begin{array}{l} 1)\ x_n > c \\ 2)\ x_n \to c, \end{array}\right\}\ (n = 1,\ 2,\ \ldots), \tag{69}$$

последовательность $\{f(x_n)\}$ стремится к одному и тому же пределу L'',

$$f(x_n) \to L'', \tag{70}$$

то говорят, что *при неограниченном приближении x к c справа функция $f(x)$ стремится к пределу L''*, и записывают:

$$f(c+0) = L''.$$

Рассмотренные нами примеры подобраны таким образом, что c всё время равно нулю. Мы получаем:

в примере 1: $L' = 0$, $L'' = +\infty$,
» » 2: $L' = 1$, $L'' = 0$,
» » 3: а) $L' = -\infty$, $L'' = +\infty$; б) $L' = L'' = +\infty$,
» примерах 4 и 5: пределы L' и L'' не существуют,
» » 6 и 7: $L' = L'' = 0$.

Следует обратить особое внимание на то, что при нахождении пределов L' и L'' число $f(c)$ — значение функции в самой рассматриваемой точке $x = c$ — не играет никакой роли: функция может и «не иметь смысла» в этой точке.

Определение предела L' допускает следующую «эпсилонную» перефразировку. Соотношение (68) означает

1) при L' конечном:

как бы мало ни было $\varepsilon\ (>0)$, можно указать такое число $\delta \equiv \delta_\varepsilon\ (>0)$, что из неравенств

$$x < c,\quad x > c - \delta \tag{71}$$

следует неравенство

$$|f(x) - L'| < \varepsilon, \tag{72}$$

2) при $L' = +\infty$:

как бы велико ни было M, можно указать такое число $\delta \equiv \delta_M\ (>0)$, что из неравенств (71) следует неравенство

$$f(x) > M, \tag{73}$$

3) при $L'' = -\infty$

— такая же формулировка, с заменой последнего неравенства следующим:

$$f(x) < -M. \tag{74}$$

Аналогичные определения могут быть сформулированы для предела L''.

Мы не будем останавливаться на доказательстве равносильности этих определений с прежним определением. Оно строится по образцу доказательства, приведённого в предыдущем параграфе.

§ 42. Двусторонний предел. Понятие непрерывности

Величины L' и L'', определённые в предыдущем параграфе, называются *левым* и *правым пределами функции* $f(x)$ в точке $x=c$. Это — *односторонние* пределы.

Если оказывается, что левый и правый пределы между собой равны

$$L'=L''=L,\ ^{1})$$

то их общее значение L является уже *двусторонним* пределом, или просто *пределом* функции $f(x)$ в точке $x=c$, чему соответствуют записи:

$$1)\ \lim_{x\to c} f(x)=L, \tag{75}$$

или

$$2)\ f(x)\to L \text{ при } x\to c. \tag{76}$$

Независимо от односторонних пределов двусторонний предел может быть определён следующим образом. Пусть функция $f(x)$ задана в некоторой окрестности точки c (исключая, может быть, самую точку c). Если для всякой последовательности точек $\{x_n\}$, взятых в этой окрестности и обладающих свойствами

$$\begin{aligned} &1)\ x_n\neq c,\\ &2)\ x_n\to c, \end{aligned} \tag{77}$$

последовательность $\{f(x_n)\}$ стремится к одному и тому же, конечному или бесконечному, пределу L,

$$f(x_n)\to L, \tag{78}$$

то говорят, что при неограниченном приближении x к c функция $f(c)$ стремится к пределу L.

«Эпсилонная» формулировка, строго эквивалентная предыдущей, такова: соотношение (75) или (76) обозначает:

1) в случае L конечного:

как бы мало ни было $\varepsilon\ (>0)$, *можно указать такое* $\delta\equiv\delta_\varepsilon\ (>0)$, *что из неравенства*

$$x\neq c,\quad |x-c|<\delta \tag{79}$$

следует неравенство

$$|f(x)-L|<\varepsilon; \tag{80}$$

2) при $L=+\infty$:

как бы велико ни было M, *можно указать такое* $\delta\equiv\delta_M\ (>0)$, *что из неравенства* (79) *следует неравенство*

$$f(x)>M; \tag{81}$$

---

[1]) Напомним, что при этом не исключаются случаи

$$L=+\infty,\ \text{или}\ L=-\infty.$$

3) при $L = -\infty$

— такая же формулировка, с заменой последнего неравенства следующим:

$$f(x) < -M. \tag{82}$$

Действительно, если соотношение (78) осуществляется для всякой последовательности $\{x_n\}$, удовлетворяющей требованию (77), то оно осуществляется и для всякой последовательности, удовлетворяющей требованию (67), или требованию (69), так что если существует предел L, то обязательно существуют пределы L' и L'', и притом $L' = L'' = L$. Обратно, пусть существуют равные между собой пределы L' и L'', и L есть их общее значение: $L' = L'' = L$. В таком случае, последовательность $\{x_n\}$, удовлетворяющую требованию (77), можно считать «составленной» [1]) из двух последовательностей, соответственно удовлетворяющих требованиям (67) и (69); тогда последовательность $\{f(x_n)\}$ будет «составлена» из двух последовательностей, имеющих один и тот же предел L; значит, $f(x_n) \to L$.

В примерах 1, 2, 3*а*, 4 и 5 предшествующего параграфа предел рассматриваемых функций в точке $x = 0$ не существует; в примерах 6 и 7 $L = \lim\limits_{x \to 0} f(x) = 0$; в примере 3б $L = +\infty$.

Чрезвычайно важно понять, что числовое значение предела, даже «двустороннего»,

$$L = \lim_{x \to 0} f(x),$$

как и самый факт его существования, не зависит от того, чему равно значение $f(c)$ функции $f(x)$ в точке $x = c$, и даже от того, определена ли сама функция в этой точке.

С понятием предела функции в данной точке тесно связано понятие непрерывности функции в этой точке.

Предположим, что функция $f(x)$, заданная в окрестности некоторой точки $x = c$, имеет конечные левый и правый пределы в этой точке, L' и L'', и что три величины L', L'' и $f(c)$ равны между собой:

$$L' = L'' = f(c). \tag{83}$$

Заметим, что по доказанному достаточно предположить, что существует конечный предел L функции $f(x)$ в точке $x = c$ и что две величины L и $f(c)$ равны между собой:

$$L = f(c). \tag{84}$$

В таком случае говорят, что функция $f(x)$ *непрерывна в точке* $x = c$.

---

[1]) См. § 37, пример 9.

Ни в одном из примеров 1—7 § 41 функция $f(x)$ не является непрерывной в точке $x=0$, и именно по следующим причинам:

в примерах 1, 2, 3а $L' \neq L''$, значит, L не существует,

в примере 3б L существует, но не конечно,

в примерах 4 и 5 L' и L'' не существуют,

в примерах 6 и 7 L существует и конечно, но функция не задана в точке $x=0$.

Впрочем, все приведённые примеры имеют то общее, что ни в одном из них функция не задана в рассматриваемой точке: этого обстоятельства достаточно, чтобы функция в ней не могла быть признана непрерывной.

В приведённом выше определении непрерывности функции $f(x)$ в данной точке c используется понятие предела: предел функции в рассматриваемой точке равен её значению в этой точке. Можно, однако, избегнуть упоминания о пределе, раскрывая содержание этого понятия в самой формулировке определения непрерывности.

Так мы приходим к определениям, которые строго эквивалентны сформулированным выше:

1. Функция $f(x)$, заданная в некоторой окрестности точки $x=c$, *непрерывна в этой точке*, если, какова бы ни была последовательность $\{x_n\}$, обладающая свойством

$$x_n \to c, \text{[1]} \tag{85}$$

непременно последовательность $\{f(x_n)\}$ сходится к пределу

$$f(x_n) \to f(c) \tag{86}$$

(определение Гейне).

2. Функция $f(x)$, заданная в некоторой окрестности точки $x=c$, непрерывна в этой точке, если, как бы мало ни было число $\varepsilon\,(>0)$, можно указать такое $\delta \equiv \delta_\varepsilon$, что из неравенства

$$|x-c| < \delta \text{[2]} \tag{87}$$

следует неравенство

$$|f(x)-f(c)| < \varepsilon \tag{88}$$

(«эпсилонное» определение Коши).

С понятием непрерывности функции в данной точке связывается понятие непрерывности функции в данном промежутке.

Говорят, что функция *непрерывна в некотором промежутке* [3], если она непрерывна в каждой точке этого промежутка.

Например, функция $f(x)=\frac{1}{x}$ непрерывна в промежутке $0<x\leqslant 1$ (если не включать начальной точки $x=0$), но она не является непрерывной в промежутке $0\leqslant x\leqslant 1$ (если включена точка $x=0$), так как не задана в точке $x=0$.

---

[1]) Необходимость ограничения $x_n \neq c$ здесь уже отпадает.

[2]) Ограничение $x \neq c$ излишне.

[3]) Промежуток может здесь включать или не включать начальную и конечную точки, или включать одну из них.

Понятие «разрыва в точке $x=c$» уточняется теперь следующим образом: если функция определена (задана) в некотором промежутке, внутри которого находится данная точка $x=c$, но не определена или не непрерывна в самой этой точке, то говорят, что в этой точке функция «имеет разрыв».

§ 43. Примеры непрерывных функций

Установить непрерывность данной функции $f(x)$ в данной точке c, по большей части, можно, не прибегая к исследованию всевозможных последовательностей вида $\{f(x_n)\}$, где $x_n \to c$ (что практически и неосуществимо), и не пользуясь также «эпсилонными» построениями.

Заметим предварительно, что функции, сводящиеся к постоянным

$$f(x) \equiv C,$$

а также функция, тождественно равная независимой переменной,

$$f(x) \equiv x,$$

очевидно, являются непрерывными в любой точке. Действительно, в первом случае мы видим (обращаясь хотя бы к «эпсилонному» определению), что отклонение $|f(x)-f(c)|$ при всяком x равно нулю, во втором — оно всегда равно отклонению $|x-c|$, и потому δ можно всегда взять равным ε.

В дальнейшем же открываются обширные возможности, опираясь на основные теоремы о пределах, из свойства непрерывности одних функций заключать о непрерывности других.

Справедливы прежде всего такие теоремы: если каждая из функций $u(x)$ и $v(x)$ непрерывна в точке c, то непрерывны в этой точке также их a) сумма $u(x)+v(x)$, b) разность $u(x)-v(x)$, c) произведение $u(x)v(x)$, d) отношение $\frac{u(x)}{v(x)}$ (при дополнительном условии $v(c) \neq 0$).

Докажем в качестве примера пункт a). По предположению из $x_n \to c$ следует $u(x_n) \to u(c)$ и $v(x_n) \to v(c)$. Тогда по теореме 1 § 38 мы получаем: $u(x_n)+v(x_n) \to u(c)+v(c)$, что и требовалось доказать.

Установим, наконец, следующую общую теорему e) о непрерывности «сложной» функции: *если функция $u(x)$ непрерывна в точке $x=c$, и функция $f(u)$ непрерывна в точке $u=u(c)$, то функция $F(x) \equiv f(u(x))$ непрерывна в точке $x=c$.*

В самом деле, если $x_n \to c$, то по свойству непрерывности функции $u(x)$

$$u(x_n) \to u(c),$$

и тогда по свойству непрерывности $f(u)$

$$f(u(x_n)) \to f(u(c)),$$

т. е. $F(x_n) \to F(c)$.

Из теоремы e) следует: если функция $v(x)$ непрерывна в промежутке $a<x<b$, если все значения, которые она принимает в этом промежутке, принадлежат промежутку $c<x<d$, и если, наконец, функция $u(x)$ определена и непрерывна в промежутке $c\leqslant x\leqslant d$, то сложная функция $u(v(x))$ непрерывна в промежутке $a<x<b$.

В частности, если функции $u(x)$ и $v(x)$ определены и непрерывны при всех действительных значениях x, то можно сказать то же самое о сложной функции $u(v(x))$.

Принимая во внимание определение целой и дробной рациональной функции (см. § 6), мы выводим, опираясь на пункты a) — d), в частности, такие заключения:

Всякий рациональный многочлен (целая рациональная функция) $P(x)$ *непрерывен всюду, т. е. в любой точке* (другими словами, в промежутке $-\infty<x<+\infty$).

Всякая дробная рациональная функция

$$R(x)=\frac{P(x)}{Q(x)}$$

(где $P(x)$ и $Q(x)$ — многочлены без общих множителей) *непрерывна всюду, кроме точек, в которых знаменатель* $Q(x)$ *обращается в нуль.* Пользуясь обозначениями предыдущего параграфа, можно сказать, что в таких точках пределы L' и L'' бесконечны, притом — одного или различных знаков в зависимости от кратности нуля $Q(x)$ (см. § 16).

Всякая рациональная функция может иметь лишь конечное число «разрывов», и все они — указанного здесь типа.

Корни целых степеней

$$f(x)=\sqrt[n]{x}$$

являются непрерывными функциями всюду, где они существуют.

Докажем это хотя бы для случая арифметического квадратного корня ($n=2$). Пусть $c>0$. Тогда

$$|\sqrt{x_n}-\sqrt{c}|=\left|\frac{x_n-c}{\sqrt{x_n}+\sqrt{c}}\right|=\frac{|x_n-c|}{\sqrt{x_n}+\sqrt{c}}\leqslant\frac{|x_n-c|}{\sqrt{c}}.$$

Если $x_n\to c$, то правая часть стремится к нулю; значит, и левая. В случае $c=0$ речь может итти лишь об «односторонней» непрерывности функции $f(x)=\sqrt{x}$ (т. е. при ограничении $x\geqslant 0$). Нужно лишь доказать, что $\sqrt{x_n}\to 0$, если $x_n>0$ и $x_n\to 0$. Чтобы из неравенства

$$|x_n|<\delta$$

следовало неравенство

$$\sqrt{x_n}<\varepsilon,$$

достаточно положить $\delta<\delta_\varepsilon=\varepsilon^2$.

По поводу общего случая ($n\geqslant 3$) см. стр. 246, пример 2.

Тригонометрические функции $\sin x$ и $\cos x$ *всюду непрерывны.*

В самом деле, возьмём на единичном круге $\smile AC = c$, $\smile AM = x$ $(x \neq c)$ и обозначим через C_1 и M_1 проекции точек C и M на горизонтальную ось, а через P — проекцию точки C на M_1M; тогда получим:

$$\sin x - \sin c = M_1M - C_1C = PM,$$

$$\cos x - \cos c = OM_1 - OC_1 = -PC;$$

далее мы видим, что

$$\left.\begin{array}{l} |\sin x - \sin c| = |PM| \\ |\cos x - \cos c| = |PC| \end{array}\right\} < CM < \smile CM = |x - c|,$$

отсюда понятно, что если $x_n \to c$, то $\sin x_n \to \sin c$ и $\cos x_n \to \cos c$.

Тригонометрические многочлены (целые рациональные функции от аргументов $\cos x$ и $\sin x$) *всюду непрерывны* (см. § 27). Мы убеждаемся в этом, ссылаясь на непрерывность «сложной функции».

О дробных тригонометрических функциях можно сказать то же, что и о дробных рациональных — с тем отличием, что если имеются разрывы функции (нули знаменателя), то число их не может быть конечным; но в пределах периода число разрывов конечно. Так, элементарные тригонометрические функции $\operatorname{tg} x$, $\sec x$ и $\operatorname{cosec} x$ имеют по два разрыва в периоде.

Как произведение непрерывных функций всюду непрерывна функция $x \sin x$ и непрерывны вообще функции вида $P(x) \sin x$ или $P(x) \cos x$, где $P(x)$ — рациональный многочлен.

Функция $f(x) = \dfrac{\sin x}{x}$ как отношение непрерывных функций, непрерывна всюду, кроме точки $x = 0$. Что касается точки $x = 0$, то в ней функция $f(x)$ не определена, и потому о непрерывности в ней не может быть речи [1]). Тем не менее, существует двусторонний предел функции $f(x)$ в точке $x = 0$, именно:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1. \tag{89}$$

Пусть сначала $x > 0$. Тогда при условии $x < \dfrac{\pi}{2}$ справедливы неравенства (§§ 25 и 29)

$$\sin x < x < \operatorname{tg} x. \tag{90}$$

Разделив все члены неравенства (90) на $\sin x$ и перейдя к обратным величинам, получим:

$$\cos x < \frac{\sin x}{x} < 1. \tag{91}$$

Вследствие непрерывности $\cos x$ в точке $x = 0$, как бы мало ни было ε, можно указать такое δ, что при $|x| < \delta$

$$|\cos x - 1| < \varepsilon,$$

[1]) См., впрочем, стр. 504, упражнение 5.

так что

$$\cos x > 1 - \varepsilon.$$

Но тогда, как следует из (91),

$$1 - \varepsilon < \frac{\sin x}{x} < 1,$$

и значит,

$$\left| \frac{\sin x}{x} - 1 \right| < \varepsilon.$$

Таким образом, в данном случае существует правый предел $L'' = 1$; существование левого предела L' и равенство $L' = L'' = 1$ следуют из чётности функции $\frac{\sin x}{x}$.

Особого внимания заслуживает вопрос о непрерывности показательной функции.

Показательная функция $f(x) = a^x$ непрерывна всюду.

Убедимся в этом, ограничиваясь случаем $a > 1$ [1]): если $a < 1$, то достаточно будет указать, что $a^x = \frac{1}{(a^{-1})^x}$ и $a^{-1} > 1$. Мы установим сначала, что существует правый предел L'' функции a^x в точке $x = c$ и притом

$$L'' = a^c. \tag{92}$$

Пусть $x_n > c$, $x_n \to c$; тогда

$$a^{x_n} = a^c \cdot a^{x_n - c}. \tag{93}$$

Положим $x_n = c + h_n > 0$. Нам достаточно показать, что из соотношения

$$h_n \to 0 \tag{94}$$

(какова бы ни была последовательность положительных чисел $\{h_n\}$) следует соотношение

$$a^{h_n} \to 1.$$

В таком случае правая часть, а значит, и левая, в равенстве (93) имеют предел a^c.

Рассмотрим последовательно несколько случаев:

1) Допустим, что $h_n = \frac{1}{n}$, и докажем, что

$$\sqrt[n]{a} \to 1. \tag{95}$$

На стр. 81 мы имели неравенство (при $a > 1$)

$$a^n > 1 + (a - 1)n.$$

Придадим ему вид

$$a - 1 < \frac{a^n - 1}{n}.$$

---

[1]) Следующее ниже доказательство одинаково применимо к случаю, когда рассматриваются лишь рациональные значения x, и к более общему случаю, когда рассматриваются какие угодно действительные значения x. Что касается определения показательной функции a^x при иррациональных значениях x, а также соответствующего распространения её свойств, то по этому поводу см. § 51.

Так как здесь a — произвольное число, с одним лишь ограничением $a>1$, то можно заменить a через $\sqrt[n]{a}$:

$$\sqrt[n]{a}-1<\frac{a-1}{n}.$$

Левая часть последнего неравенства положительна, и так как правая, очевидно, стремится к нулю, то следовательно, то же справедливо и относительно левой части. Отсюда вытекает (95).

2) Пусть $h_n=\frac{1}{p_n}$, где члены последовательности $\{p_n\}$ — целые положительные числа, причём $p_n\to\infty$. Тогда соотношение

$$\sqrt[p_n]{a}\to 1$$

является прямым следствием из соотношения (95) (см. стр. 168—169).

3) Пусть члены последовательности $\{h_n\}$ — какие угодно положительные числа, $h_n\to 0$. Подберём целые числа p_n по условию

$$p_n\leqslant\frac{1}{h_n}<p_n+1 \qquad (n=1,\ 2\ \ldots);$$

очевидно, $p_n\to\infty$. Тогда

$$h_n\leqslant\frac{1}{p_n},$$

и так как функция a^x — возрастающая, то

$$a^{h_n}\leqslant a^{\frac{1}{p_n}}.$$

Левая часть в последнем неравенстве больше 1, правая, согласно предыдущему пункту, имеет предел 1; значит, $a^{h_n}\to 1$.

Существование левого предела L' функции a^x в точке c и равенство

$$L'=a^c \tag{96}$$

можно доказать, исходя из тождества

$$a^{x_n}=\frac{a^c}{a^{c-x_n}} \qquad (x_n<c)$$

и полагая

$$c-x_n=h_n>0.$$

Из соотношений (92) и (96), принимая во внимание, что a^c есть значение функции a^x в точке c, следует непрерывность функции a^x в этой точке.

Функция $\log_a x\,(a>1)$ *непрерывна во всех точках, где она существует, т. е. при* $x>0$.

Это вытекает из свойства возрастания показательной функции a^x. Рассмотрим точку $x=c>0$; покажем, что правый предел $\log_a x$ в этой точке, L'', существует и равен $\log_a c$. Пусть $x_n>c$, $x_n\to c$. Из $x_n>c$ следует, что $\log_a x_n>\log_a c$. Нужно доказать, что $\log_a x_n\to\log_a c$. Допустим, что это неверно. В таком случае существует такое число $\varepsilon^*\,(>0)$, что для сколь угодно больших зна-

чений n имеет место неравенство

$$\log_a x_n > \log_a c + \varepsilon^*.$$

Но тогда для тех же значений n (по свойству возрастания a^x)

$$a^{\log_a x_n} > a^{\log_a c + \varepsilon^*},$$

т. е.

$$x_n > ca^{\varepsilon^*} > c.$$

А это противоречит допущению $x_n \to c$.

Итак, доказано существование правого предела L'' в точке c и равенство $L'' = \log_a c$. Подобным же образом доказывается существование в этой точке левого предела L' и равенство $L' = \log_a c$. Отсюда же следует равенство $L' = L'' = \log_a c$ и, следовательно, непрерывность функции $\log_a x$ в точке c [1]).

Непрерывность общей степенной функции x^α при любом постоянном α, в любой точке x ($x > 0$) получается как следствие теоремы о непрерывности сложной функции. Действительно (§ 24),

$$x^\alpha = a^{\alpha \log_a x},$$

и, полагая $v(x) = \alpha \log_a x$, $u(x) = a^x$, мы видим, что функция $v(x)$ непрерывна при $x > 0$, функция же $u(x)$ непрерывна при любом значении x, откуда следует заключение.

Таким же образом устанавливается непрерывность функций вроде x^x при $x > 0$ ($x^x = a^{x \log_a x}$).

Подводя некоторые итоги сказанному, позволительно констатировать что для элементарных функций точки непрерывности являются, так сказать правилом, а точки разрыва — исключением. Многие элементарные функции непрерывны во всех точках, где они заданы; другие непрерывны всюду, кроме «отдельных» точек (так часто говорили раньше, не видя необходимости в разъяснении того, что такое «отдельные точки»).

Приведём в качестве заключительного примера функцию

$$f(x) = \frac{1}{\sin \frac{1}{x}}, \tag{97}$$

имеющую разрывы во всех точках вида $x = \frac{1}{k\pi}$, где k — целые числа ($k \neq 0$), и, кроме того, ещё разрыв в точке $x = 0$.

§ 44. Пределы при монотонном изменении. Число e

Теорема Вейерштрасса. *Всякая возрастающая, или хотя бы неубывающая, ограниченная сверху последовательность $\{a_n\}$ имеет конечный предел.*

При этом под *возрастающей* последовательностью, разумеется, нужно понимать такую, у которой каждый следующий член больше

[1]) Более общая теорема по поводу непрерывности обратных функций приведена в § 52.

предыдущего; если же говорят о *неубывающей* последовательности, то предполагают, что каждый следующий член может быть больше или равен предыдущему:

$$a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots \tag{98}$$

Раз последовательность — возрастающая или неубывающая, то она непременно ограничена снизу, так как все её члены больше некоторого числа m, меньшего, чем первый член. Так как по условию она, кроме того, ограничена и сверху, то её можно назвать просто ограниченной. Но в таком случае по теореме Больцано-Вейерштрасса (§ 36) она имеет хотя бы одну предельную точку: нужно доказать, что такая точка — только одна. Допущение, что имеется более одной предельной точки, приведёт к противоречию. В самом деле, пусть последовательность $\{a_n\}$ в числе своих предельных точек имеет две точки a и a', и пусть, например, $a < a'$. Выберем число ε по условию

$$0 < \varepsilon < \frac{a' - a}{2};$$

тогда

$$a + \varepsilon < a' - \varepsilon. \tag{99}$$

В ε-окрестности точки a' должна находиться хоть одна точка последовательности $\{a_n\}$, например a_N:

$$a' - \varepsilon < a_N < a' + \varepsilon.$$

В таком случае все члены последовательности a_n, следующие за a_N, будут не меньше чем $a' - \varepsilon$:

$$a_n \geqslant a_N > a' - \varepsilon \qquad (n = N + 1,\ N + 2,\ \ldots)$$

и значит, меньшими чем $a' - \varepsilon$ смогут оказаться лишь члены, имеющие индекс n, меньший чем N: таких членов — конечное число. Итак, в ε-окрестности точки a вследствие (99) сможет заключаться лишь конечное число членов. А это противоречит тому, что a есть предельная точка.

Таким образом, последовательность $\{a_n\}$ имеет лишь одну предельную точку, и она есть её предел: последовательность — сходящаяся.

Заметим, что в условии теоремы Вейерштрасса нет необходимости предполагать, что последовательность — неубывающая с первого же члена: она может быть неубывающей, начиная с *некоторого* члена.

То, что справедливо относительно неубывающей, ограниченной сверху, последовательности, переносится, конечно, и на невозрастающую, ограниченную снизу последовательность.

Условившись (как иногда делают) называть *монотонной*[1]) неубывающую или невозрастающую последовательность, можно даже охватить оба случая общей формулировкой:

Всякая монотонная ограниченная последовательность имеет конечный предел.

Или, короче:

Всякая монотонная последовательность имеет предел. Если последовательность ограничена, предел — конечный, если не ограничена, — бесконечный.

Пример 1. Возрастающая последовательность

$$\left\{1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{n^2}\right\}\equiv 1,\quad 1+\frac{1}{2^2},\quad 1+\frac{1}{2^2}+\frac{1}{3^2},\ldots \tag{100}$$

ограничена сверху.

Действительно, по поводу члена a_{2^p-1} можно утверждать:

$$a_{2^p-1}=1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{(2^p-1)^2}=$$

$$=\left\{\begin{array}{l}1+\\ +\frac{1}{2^2}+\frac{1}{3^2}+\\ +\frac{1}{4^2}+\frac{1}{5^2}+\frac{1}{6^2}+\frac{1}{7^2}+\\ +\ldots\ldots\ldots\ldots+\\ +\left(\frac{1}{2^{p-1}}\right)^2+\ldots+\left(\frac{1}{2^p-1}\right)^2\end{array}\right\}<\left\{\begin{array}{l}1+\\ +\frac{1}{2^2}+\frac{1}{2^2}+\\ +\frac{1}{4^2}+\frac{1}{4^2}+\frac{1}{4^2}+\frac{1}{4^2}+\\ +\ldots\ldots\ldots\ldots+\\ +\left(\frac{1}{2^{p-1}}\right)^2+\ldots+\left(\frac{1}{2^{p-1}}\right)^2\end{array}\right\}<$$

$$<1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^{p-1}}\ ^{2})=2-\frac{1}{2^{p-1}}<2.$$

Раз неравенство

$$a_{2^p-1}<2 \tag{101}$$

доказано при любом p, то справедливо также при любом n неравенство

$$a_n<2. \tag{102}$$

В самом деле, к заданному n всегда можно подобрать p так, чтобы было

$$2^p-1\geqslant n;$$

тогда будем иметь:

$$a_n\leqslant a_{2^p-1}$$

и, значит, с помощью (101) получится неравенство (102).

Итак, существует предел

$$a=\lim a_n=\lim\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{n^2}\right)\leqslant 2.$$

---

[1]) См. § 5, стр. 37.

[2]) Произведено суммирование по строкам.

С помощью разного рода приёмов не особенно трудно вычислить этот предел с наперёд назначенным числом десятичных знаков:

$$a = 1{,}644\ldots$$

Но лишь математический анализ позволяет установить, что этот предел замечательным образом связан с числом π:

$$a = \frac{\pi^2}{6}$$

(последний результат принадлежит Эйлеру).

Пример 2. Последовательность $\{a_n\}$, где

$$a_n = \left(1 + \frac{1}{n}\right)^n, \tag{103}$$

имеет особенно большое значение в анализе.

Можно порекомендовать читателю вычислить (непосредственно) значения a_n при $n = 1, 2, 3, 4, 5$, округляя их до 0,001; затем с помощью логарифмов при $n = 10$, 50, 100 и даже $n = 1000$.

Последовательность (103) — возрастающая. В самом деле, раскрывая по формуле бинома Ньютона, мы получаем:

$$a_n = 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \\ + \ldots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{n-1}{n}\right) \tag{104}$$

и, таким же образом,

$$a_{n+1} = 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n+1}\right) + \frac{1}{3!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right) + \\ + \ldots + \frac{1}{n!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right)\ldots\left(1 - \frac{n-1}{n+1}\right) + \frac{1}{(n+1)^{n+1}}.$$

Каждый член во второй сумме, начиная с третьего, больше, чем соответствующий член первой суммы, и имеется ещё лишний, последний, член; значит,

$$a_n < a_{n+1}.$$

С другой стороны, последовательность $\{a_n\}$ ограничена сверху. Действительно, из формулы (104) следует, что, каково бы ни было n,

$$a_n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} < \\ < 1 + \left(1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}\right) < 1 + 2 = 3.$$

Отсюда следует, что последовательность $\{a_n\}$ имеет конечный предел, $\leqslant 3$. Этот предел обозначается буквой e:

$$e = \lim\left(1 + \frac{1}{n}\right)^n. \tag{105}$$

Его числовое значение даётся равенством

$$e = 2{,}71828\ldots$$

Число e играет большую роль в математическом анализе: его принимают в качестве основания системы «натуральных» логарифмов [1]) (см. стр. 316).

Сказанное относительно последовательностей оказывается справедливым и для функций.

Если функция $f(x)$, заданная в некоторой окрестности точки $(+\infty)$, является возрастающей (или хотя бы неубывающей) и обладает свойством ограниченности сверху, то существует конечный предел

$$L = \lim_{x \to \infty} f(x).$$

Нужно доказать, что, какова бы ни была последовательность $\{x_n\}$ в заданной окрестности точки $(+\infty)$, из условия $x_n \to \infty$ следует, что последовательность $\{f(x_n)\}$ стремится к некоторому одному и тому же пределу.

Последовательность $\{f(n)\}$ — возрастающая и ограниченная; положим

$$L = \lim f(n). \tag{106}$$

Пусть p_n — такие целые числа, что

$$p_n \leqslant x_n < p_n + 1 \qquad (n = 1, 2, \ldots). \tag{107}$$

Из неравенства (107) следует:

$$f(p_n) \leqslant f(x_n) < f(p_n + 1), \tag{108}$$

и так как, очевидно, $p_n \to \infty$ и $p_n + 1 \to \infty$, а из соотношения (106) вытекает $f(p_n) \to L$ и $f(p_n + 1) \to L$, то отсюда на основании неравенства (108) заключаем:

$$f(x_n) \to L.$$

Подобная же теорема справедлива, конечно, и в случае функции $f(x)$, *убывающей* (невозрастающей) *и ограниченной снизу* в окрестности точки $(+\infty)$.

Аналогичные формулировки существуют и для точки $(-\infty)$. Наконец, не исключается и случай одностороннего приближения неизвестной переменной к конечному пределу:

Если функция $f(x)$, заданная в некотором промежутке $c < x < c + \delta$ или $c - \delta < x < c$, где $\delta > 0$, изменяется монотонно (т. е. не убывая или не возрастая), *то существует предел $f(c + 0)$* (или $f(c - 0)$).

Этот предел конечен в том случае, если функция в данном промежутке ограничена.

---

[1]) Логарифмы, взятые по основанию e, часто обозначаются следующим образом:

$$\log_e x = \ln x.$$

Нет необходимости приводить доказательство этого утверждения.

Пример 3. Если бы было легко доказать элементарными средствами, что функция

$$f(x)=\left(1+\frac{1}{x}\right)^x$$

— возрастающая, то она могла бы хорошо иллюстрировать предыдущую теорему. Не имея, однако, возможности опираться на это свойство, мы, чтобы доказать равенство

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e, \tag{109}$$

станем рассуждать несколько иначе. Нам уже известно (см. при-

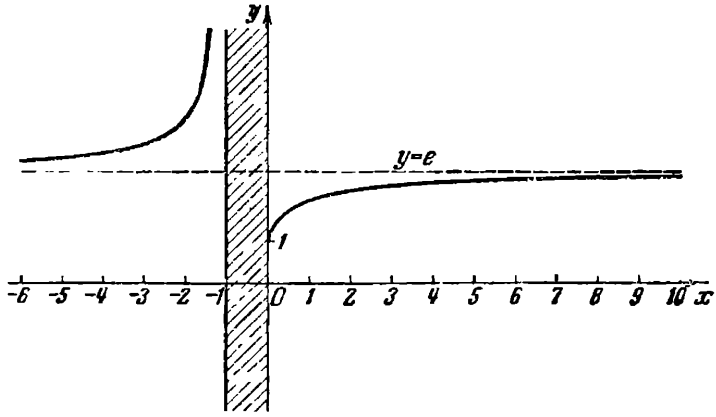


Рис. 78.

мер 2), что $f(n)\to e$; предположим теперь, что $x_n\to\infty$, и докажем, что $f(x_n)\to e$. Пусть (как раньше)

$$p_n\leqslant x_n<p_n+1,$$

где p_n — целое. Тогда

$$\left(1+\frac{1}{x_n}\right)^{x_n}<\left(1+\frac{1}{p_n}\right)^{p_n+1}=\left(1+\frac{1}{p_n}\right)^{p_n}\cdot\left(1+\frac{1}{p_n}\right) \tag{110}$$

и, с другой стороны,

$$\left(1+\frac{1}{x_n}\right)^{x_n}>\left(1+\frac{1}{p_n+1}\right)^{p_n}=\frac{\left(1+\frac{1}{p_n+1}\right)^{p_n+1}}{1+\frac{1}{p_n+1}}. \tag{111}$$

Нетрудно сообразить, что правые части в обоих неравенствах (110) и (111) имеют предел e; отсюда следует соотношение

$$\lim\left(1+\frac{1}{x_n}\right)^{x_n}=e$$

и, наконец, (109).

Рассмотрим также, чему равняется $\lim\limits_{x \to -\infty} f(x)$. Пусть $\{x_n\}$ — такая последовательность значений x, что $x_n \to -\infty$. Положим $|x_n| = -x_n = \xi_n$. Тогда

$$f(x_n) = \left(1 + \frac{1}{x_n}\right)^{x_n} = \left(1 - \frac{1}{\xi_n}\right)^{-\xi_n} = \left(1 + \frac{1}{\xi_n - 1}\right) \cdot \left(1 + \frac{1}{\xi_n - 1}\right)^{\xi_n - 1},$$

и так как $\xi_n \to \infty$ и также $\xi_n - 1 \to \infty$, то первый множитель справа стремится к 1, второй, по доказанному, — к числу e. Итак, наряду с соотношением (109) мы имеем:

$$\lim_{x \to -\infty} \left(1 + \frac{1}{x}\right)^x = e. \tag{112}$$

На рис. 78 дан график функции $y = \left(1 + \frac{1}{x}\right)^x$ [1].

---

[1]) В заштрихованной полосе могут встретиться отдельные точки данного графика. Так, при $x = -\frac{1}{2}$ не получается никакого (действительного) значения функции $f(x)$; при $x = -\frac{1}{3}$ получается одно отрицательное значение; при $x = -\frac{2}{3}$ — одно положительное значение. Осмыслить эти явления можно с помощью теории функций комплексного переменного.

ГЛАВА IV

ПРЕДЕЛЫ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ФУНКЦИЙ. СВОЙСТВА НЕПРЕРЫВНЫХ ФУНКЦИЙ

§ 45. Простая сходимость

Говоря о последовательностях чисел $\{a_n\}$, мы установили (§ 38) взгляд на переход к пределу ($a_n \to a$) как на операцию, позволяющую иногда по данным числам a_n ($n = 1, 2, 3, \ldots$) построить новое число a, называемое *пределом* последовательности:

$$a = \lim a_n.$$

Переходя теперь к рассмотрению последовательностей функций

$$\{f_n(x)\} \equiv f_1(x), \quad f_2(x), \ldots, \quad f_n(x), \ldots, \tag{1}$$

заданных в некотором (конечном или бесконечном) промежутке, определим и в этом случае операцию перехода к пределу

$$f_n(x) \to f(x),$$

позволяющую иногда по данным функциям $f_n(x)$ ($n = 1, 2, 3, \ldots$) построить новую функцию — предел последовательности данных функций

$$f(x) = \lim_{n \to \infty} f_n(x). \tag{2}$$

К вопросу о сходимости последовательности функций можно подходить с различных точек зрения. Мы установим сначала более общее понятие *простой* сходимости и лишь позже (см. § 48) обратимся к более специальному (и практически более важному) понятию *равномерной* сходимости.

Можно рассуждать так. Задавая в некотором промежутке I последовательность функций (1), мы задаём бесконечное множество числовых последовательностей: именно, если x_0 есть какое-нибудь значение x из промежутка I, то с каждым таким значением сопоставляется числовая последовательность

$$\{f_n(x_0)\} \equiv f_1(x_0), \quad f_2(x_0), \ldots, \quad f_n(x_0), \ldots \tag{3}$$

Может случиться, что некоторые из таких числовых последовательностей $\{f_n(x_0)\}$ имеют предел, и притом конечный; этот предел, вообще говоря, зависит от выбора точки x_0; обозначим его через $f(x_0)$.

Сказать, что в промежутке I последовательность $\{f_n(x)\}$ стремится к пределу $f(x)$

$$f_n(x) \to f(x), \tag{4}$$

означает, по определению, то же самое, что констатировать наличие предельного отношения (в собственном смысле — см. § 38)

$$f_n(x_0) \to f(x_0)$$

для каждой точки x_0 из промежутка I.

Нужно сразу же заметить, что символ $f(x)$ вовсе не обязательно обозначает элементарную функцию (см. § 1). В самом деле, на основании предыдущего с каждым значением x из промежутка I с о п о с т а в л я е т с я по особому правилу некоторое число $y=f(x)$; но упомянутое правило подразумевает выполнение, кроме элементарных операций, ещё операции перехода к пределу (для случая числовой последовательности), и вовсе ни откуда не следует, чтобы тот же результат мог быть получен без этой операции. Таким образом, если предел последовательности (1) мы станем трактовать как функцию числового значения переменной x, то следует заранее считать не исключённым, что функция эта уже не является элементарной, а есть функция в более общем, расширенном смысле слова. Дальнейшие примеры покажут характер получающихся этим способом функций.

Если не делать предположения, что при любом числовом значении $x=x_0$ из рассматриваемого промежутка I последовательность (1) стремится к конечному пределу, то в зависимости от выбранного значения x_0 возможны три различных случая:

1) или эта последовательность имеет конечный предел,
2) или её предел бесконечен ($+\infty$ или $-\infty$),
3) или она не имеет предела.

Обозначим соответствующие совокупности (множества) точек $x=x_0$ через E_1, E_2 и E_3; каждая точка из I принадлежит одной и только одной из этих совокупностей. Символ $f(x_0)$ имеет смысл в том случае, если точка x_0 принадлежит совокупности E_1; эта совокупность носит название *области сходимости* последовательности (1) (в пределах промежутка I). Если x_0 принадлежит E_3, то символ $f(x_0)$ не имеет никакого смысла; если же x_0 принадлежит E_2, то возможна была бы запись

$$f(x_0)=+\infty \quad \text{или} \quad f(x_0)=-\infty,$$

но она не принята [1]). Итак, в общем случае функция $f(x)$ оказывается з а д а н н о й лишь в точках совокупности E_1.

---

[1]) Тем не менее случаи, когда x_0 принадлежит E_2 или принадлежит E_3, конечно, существенно различны.

В дальнейшем запись

$$f_n(x) \to f(x) \quad \text{или} \quad \lim_{n\to\infty} f_n(x) = f(x)$$

будет использована только для случая $E_1 \equiv I$, т. е. при допущении, что сходимость к конечному пределу имеет место во всех точках рассматриваемого промежутка.

Рассмотрим несколько примеров, в которых роль промежутка I будет играть вся действительная ось $-\infty < x < +\infty$.

Пример 1. Пусть дана последовательность функций $\{f_n(x)\}$, где

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n \qquad (n = 1,\ 2,\ 3, \ldots).$$

Рассмотрим три случая:

1° $x = 0$. В этом случае, очевидно, $f_n(x) \to 1$.

2° $x > 0$. Считая x постоянным, положим

$$\frac{n}{x} = \omega_n;$$

так как $\omega_n \to \infty$, то на основании § 44 (пример 3) и в силу непрерывности степенной функции

$$f_n(x) = \left\{\left(1 + \frac{x}{n}\right)^{\frac{n}{x}}\right\}^x = \left\{\left(1 + \frac{1}{\omega_n}\right)^{\omega_n}\right\}^x \to e^x.$$

3° $x < 0$. При этом предположении мы получаем, сделав ту же подстановку, $\omega_n \to -\infty$, и потому снова (см. конец § 44)

$$f_n(x) \to e^x.$$

Итак, во всех случаях мы имеем:

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x. \tag{5}$$

В этом примере сходимость имеется при всех значениях x, и функция, получающаяся при переходе к пределу, — элементарная.

Пример 2. Положим $f_n(x) = x^n$ $(n = 1,\ 2,\ 3, \ldots)$. Последовательность

$$\{x^n\} \equiv x,\ x^2,\ x^3, \ldots,\ x^n, \ldots$$

(см. § 37, п. 5 и § 39, п. 5) сходится к пределу 0 при условии $-1 < x < 1$, сходится к пределу 1 при $x = 1$, расходится к $+\infty$ при $x > 1$ и вовсе не имеет предела при $x \leqslant -1$. Таким образом, можно сказать, что в данном случае совокупность точек E_1 (область сходимости) есть промежуток от -1 до $+1$, со включением правого конца, но без включения левого; совокупность E_2 — промежуток от 1 до ∞, без включения левого конца; совокупность E_3 — промежуток от $-\infty$ до -1, со включением правого конца. Предельная функция $f(x) = \lim\limits_{n\to\infty} x^n$ «определена», «задана», «суще-

ствуёт» при условии $-1<x\leqslant 1$, и именно определяется равенствами

$$f(x)=\begin{cases}0, & \text{если } -1<x<+1,\\ 1, & \text{если } x=1.\end{cases} \tag{6}$$

Её график, с разрывом на правом конце, изображён на рис. 79

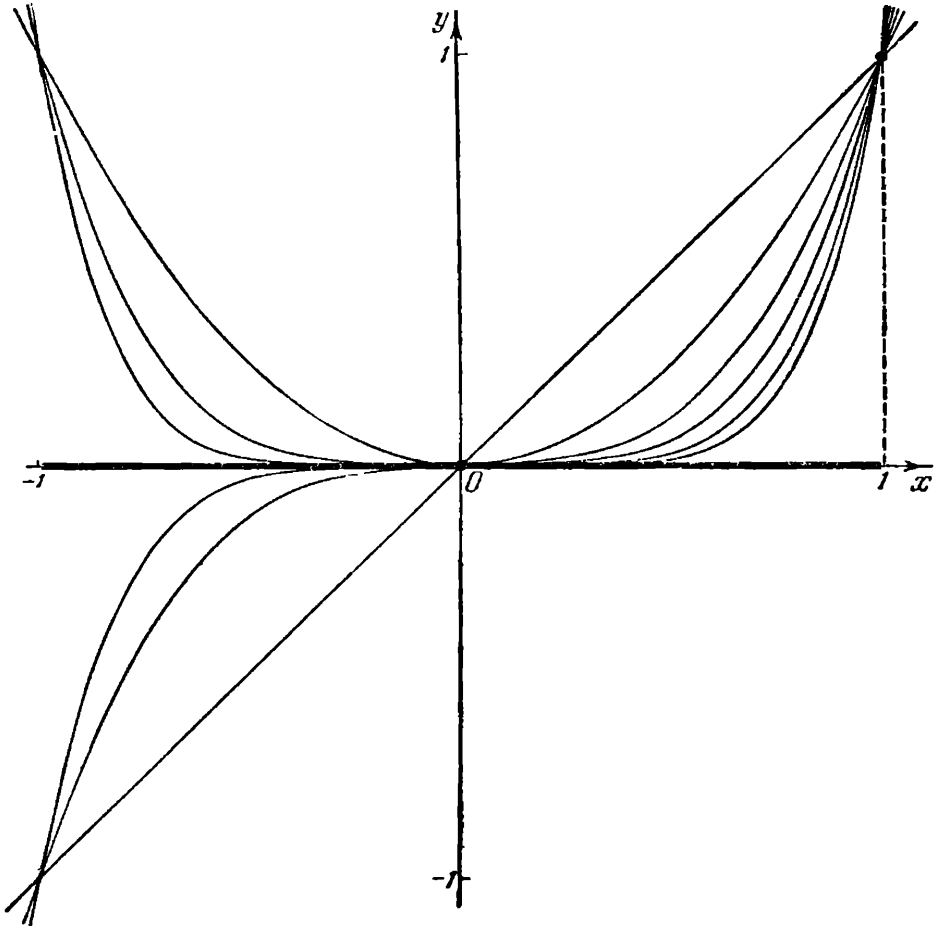


Рис. 79.

жирной линией, причём следует добавить, что точки (1, 0) и (— 1, 0) ему не принадлежат.

Пример 3. Пусть $f_n(x)=\frac{1}{1+n^2x^2}$ $(n=1,\ 2,\ 3,\ldots)$. Последовательность

$$\left\{\frac{1}{1+n^2x^2}\right\}=\frac{1}{1+x^2},\quad \frac{1}{1+4x^2},\quad \frac{1}{1+9x^2},\ldots$$

— сходящаяся при всех значениях x. Предельная функция, как легко понять, определяется равенствами

$$f(x)=\begin{cases}1 & \text{при } x=0,\\ 0 & \text{при } x\neq 0.\end{cases} \tag{7}$$

Она имеет разрыв в точке $x=0$ (рис. 80). В этой точке (см. § 41)

$$L'=f(0-0)=0, \quad L''=f(0+0)=0,$$

но

$$f(0)=1.$$

В этом примере график $f_n(x)$ получается из графика $f_1(x)$ посредством сжатия в n раз по направлению Ox.

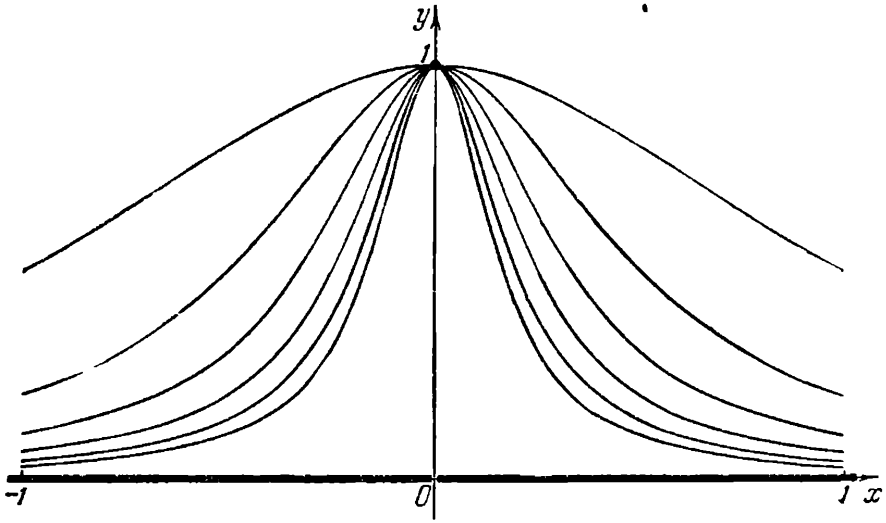


Рис. 80.

Подобные же результаты получаются, если положить

$$f_n(x)=\left(\frac{1}{1+x^2}\right)^n.$$

Пример 4.

$$f_n(x)=\cos^{2n}\pi x \qquad (n=1,\ 2,\ 3,\ldots).$$

И в этом примере имеется сходимость на всей оси; предельная функция $f(x)$ задаётся равенствами

$$f(x)=\begin{cases}1, & \text{если } x \text{ — целое},\\ 0 & \text{в других случаях.}\end{cases} \tag{8}$$

Она имеет бесконечное множество разрывов в точках, соответствующих целым значениям x (рис. 81).

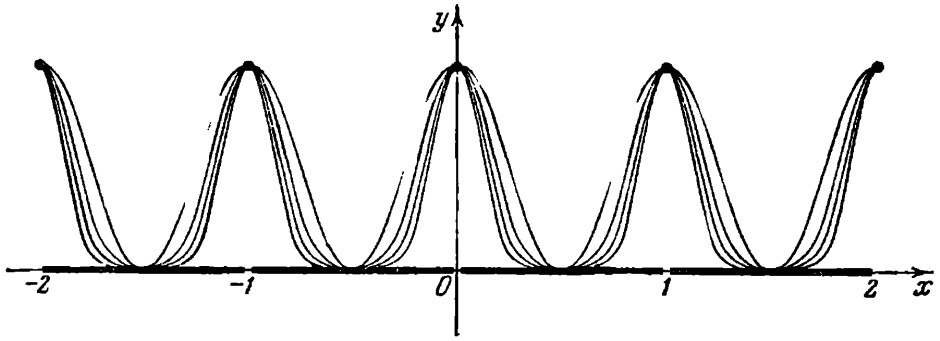


Рис. 81.

Пример 5.

$$f_n(x) = \cos^n \pi x \qquad (n = 1,\ 2,\ 3, \ldots).$$

В этом примере точки вида $x = k$, где k — целое нечётное, принадлежат множеству E_3: в них нет сходимости. Все остальные точки принадлежат множеству E_1.

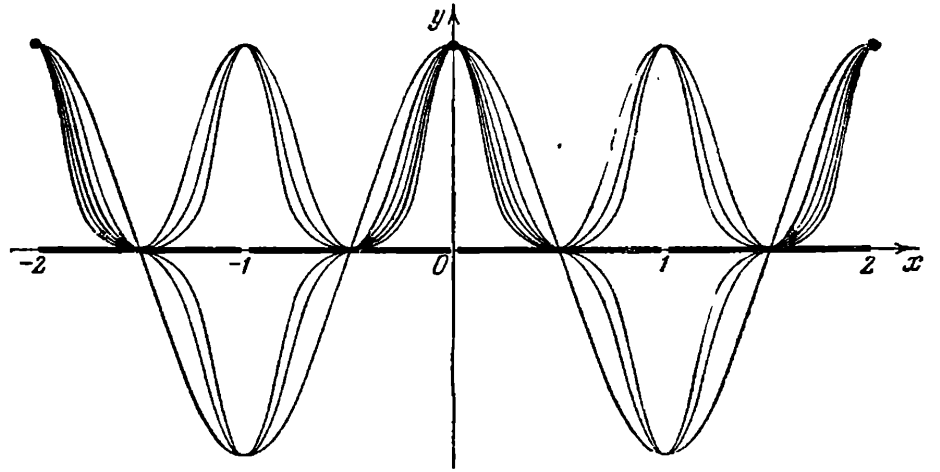


Рис. 82.

Предельная функция $f(x)$ определяется равенствами (рис. 82)

$$f(x) = \begin{cases} 1, & \text{если } x = k, \text{ где } k \text{ — целое чётное,} \\ 0, & \text{если } x \text{ — не целое число.} \end{cases} \tag{9}$$

Пример 6.

$$f_n(x) = \frac{2}{\pi} \operatorname{arctg} nx \qquad (n = 1,\ 2,\ 3, \ldots).$$

Графики функций $f_n(x)=\frac{2}{\pi}\operatorname{arctg} nx$ изображены на рис. 83. График $f_n(x)$ получается из графика $f_1(x)$ посредством сжатия в n раз по направлению оси Ox.

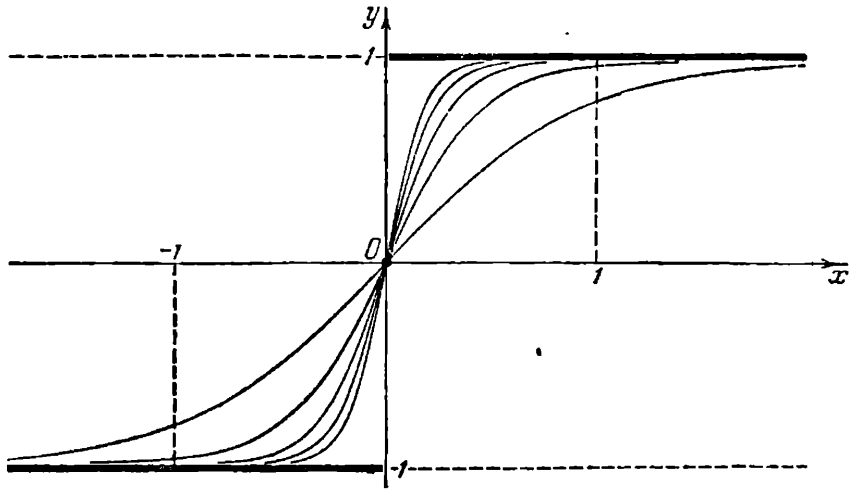


Рис. 83.

Предельная функция $f(x)$ задана для всех значений x и определяется равенствами

$$f(x)=\begin{cases} -1 \text{ при } x<0, \\ 0 \text{ при } x=0, \\ +1 \text{ при } x>0. \end{cases} \tag{10}$$

В точке $x=0$ функция $f(x)$ имеет разрыв,

$$L'=f(0-0)=-1, \quad f(0)=0, \quad L''=f(0+0)=+1.$$

Функцию $f(x)$, определённую равенствами (10), иногда обозначают $\operatorname{sgn} x$ [1]).

Пример 7.

$$f_n(x)=\frac{2}{\pi}x\operatorname{arctg} nx \qquad (n=1,\ 2,\ 3,\ldots). \tag{11}$$

Переход к пределу нам даёт:

$$f(x)=\lim_{n\to\infty} f_n(x)=\lim_{n\to\infty}\frac{2}{\pi}x\operatorname{arctg} nx=x\cdot\lim_{n\to\infty}\frac{2}{\pi}\operatorname{arctg} nx=$$
$$=x\cdot\operatorname{sgn} x=|x|.$$

[1]) sgn — сокращение латинского слова signum (знак).

График функции $f_n(x)$ из графика $f_1(x)=\frac{2}{\pi}x\operatorname{arctg}x$, изображённого на рис. 84, получается посредством одновременного сжатия в n раз по направлениям Ox и Oy (т. е. уменьшения в n раз). В пределе получается фигура, составленная из двух лучей — биссектрис 1-го и 2-го координатных углов. Предельная функция, таким образом, есть $|x|$.

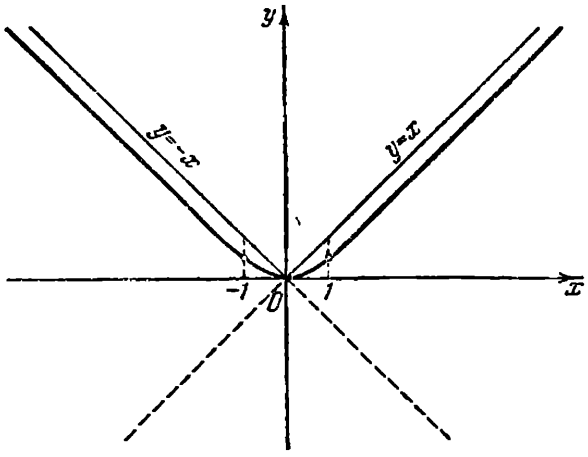


Рис. 84.

Пример 8. Последовательность $\{f(x)\}$, где

$$f_n(x)=\frac{1}{1+2^{nx}},$$

имеет предельную функцию

$$f(x)=\begin{cases}1 & \text{при } x<0,\\ \frac{1}{2} & \text{при } x=0,\\ 0 & \text{при } x>0.\end{cases} \qquad (12)$$

Пример 9. Пусть $u(x)$ и $v(x)$ — две какие угодно функции, заданные для всех значений x. Тогда последовательность $\{f_n(x)\}$, где

$$f_n(x)=u(x)\cdot\frac{1}{1+2^{nx}}+v(x)\cdot\frac{1}{1+2^{-nx}},$$

имеет предельную функцию $f(x)$, определяемую равенствами

$$f(x)=\begin{cases}u(x) & \text{при } x<0,\\ \frac{u(x)+v(x)}{2} & \text{при } x=0,\\ v(x) & \text{при } x>0.\end{cases}$$

Такой же результат получился бы, если бы мы взяли

$$f_n(x)=\frac{u(x)+v(x)}{2}-\frac{u(x)-v(x)}{2}\cdot\frac{2}{\pi}\operatorname{arctg}nx.$$

Пример 10. Пусть $F(x)$ какая угодно функция $(-\infty<x<\infty)$. Тогда последовательность $\{f_n(x)\}$, где

$$f_n(x)=\left[\frac{1}{1+F^2(x)}\right]^n,$$

имеет предельную функцию $f(x)$, равную 1 в тех точках, где $F(x)=0$, и равную 0 во всех остальных точках.

Положив подобным же образом

$$f_n(x)=\frac{2}{\pi}\operatorname{arctg} nF(x),$$

мы получим предельную функцию $f(x)=\operatorname{sgn}F(x)$, равную -1, 0 или $+1$, смотря по тому, будет ли $F(x)$ отрицательным, равным нулю или положительным числом.

Пример 11. Функция

$$f_p(x)=\lim_{q\to\infty}\cos^{2q}(10^p\pi x)$$

от предельной функции примера 4 отличается лишь тем, что график сжат в 10^p раз по направлению оси Ox. Она имеет значение 1, если x выражается десятичной дробью «ранга p» (см. § 36); во всех остальных точках она равна нулю.

Рассмотрим теперь функцию

$$f(x)=\lim_{p\to\infty}f_p(x)$$

— предел последовательности

$$f_1(x),\quad f_2(x),\quad f_3(x),\ldots$$

Эта последовательность состоит из одних нулей, если x не представляется в виде конечной десятичной дроби, и из одних единиц, — если x есть десятичная дробь ранга 1; если же, вообще говоря, x представляется в виде конечной дроби ранга $k(>1)$ (но не низшего), то первые $k-1$ членов последовательности — нули, все же остальные, начиная с k-го, — единицы. Таким образом, функция $f(x)$, т. е. предел последовательности $\{f_n(x)\}$ равна единице, если x представляется в виде конечной десятичной дроби, и равна нулю, если x таким образом не представляется.

Нарисовать график функции $y=f(x)$, конечно, невозможно, но всё же его можно «мыслить», как множество точек, «всюду плотно» расположенных на каждой из прямых $y=0$ и $y=1$ (в плоскости Oxy).

Следует обратить внимание на то, что при определении функции $f(x)$ в этом примере операция перехода к пределу выполнялась дважды:

$$f(x)=\lim_{p\to\infty}\lim_{q\to\infty}\cos^{2q}(10^p\pi x). \tag{13}$$

§ 46. Общее понятие функции одной действительной переменной

Как можно было догадываться заранее, мы заключаем, основываясь на рассмотрении предыдущих примеров, что допущение операции перехода к пределу на равных правах с прочими «элементарными» операциями открывает неограниченные возможности для расширения классов рассматриваемых функций. Так возникают функции, имеющие разрывы разнообразных типов [1])

[1]) В частности, такие, что все три числа $L'=f(c-0)$, $f(c)$ и $L''=f(c+0)$ оказываются различными между собой (см. § 41).

(см. примеры 2—6, § 45); функции, «склеенные» из нескольких данных функций (примеры 7 и 9, там же); функции, определяемые одним или другим способом в зависимости от того, какому «множеству» принадлежит значение независимой переменной (примеры 10—11, там же). Как мы видим, ограничивать себя рассмотрением лишь класса «элементарных» функций становится искусственным, плохо оправданным, запретом. К тому же легко придти к мысли, что проще и удобнее определять функцию непосредственно системой равенств вроде (7), (8) или (9), чем посредством предельных переходов, подобных указанным выше. Естественно сделать ещё один решительный шаг вперёд и сказать, что существенным в самой идее функции является то, какие значения она принимает при различных значениях переменной, а не то, какие операции нужны для вычисления её значений.

Высказанные соображения достаточны для того, чтобы не только декларировать, но и внутренне оправдать то определение понятия функции действительной переменной, которое единственно принято в настоящее время в науке.

Говорят, что на некотором числовом множестве E задана функция

$$y=f(x),$$

если с каждым значением «независимой» переменной x из множества E каким бы то ни было способом сопоставлено некоторое, только одно, значение «зависимой» переменной y.

Другими словами: указано правило, на основании которого каждому значению x из E единственным образом приводится в соответствие некоторое значение y.

Это определение необычайно расширяет понятие функции, так как характер указываемого «правила» ничем не ограничивается: в частности, роль «правила» может играть формула, содержащая элементарные операции, но это не обязательно.

Вместе с тем новое определение несколько сужает понятие функции: раз каждому значению x соответствует только одно значение $y=f(x)$, значит, тем самым устраняются из рассмотрения «многозначные функции», и остаются лишь «однозначные».

Введение понятия *однозначной* функции имеет для изложения теории функций действительного переменного первостепенное методическое значение: ограничиваясь лишь рассмотрением однозначных функций (или «расщепляя» многозначные функции на ряд однозначных), мы упрощаем формулировки многих утверждений.

Мы уже упоминали об определении функции как соответствия в § 1,— правда, лишь в предположении, что функция «задана» в некотором промежутке, а не в «произвольном множестве» точек.

Приведём ещё несколько примеров функций, задаваемых «независимо от формулы».

Пример 1. Функция, называемая «целая часть x» и обозначаемая

$$f(x)=[x], \tag{14}$$

определяется для любого значения, как наибольшее целое число, не превышающее данного значения x. Например,

$$\left[7\frac{1}{3}\right]=7, \quad [3]=3, \quad [-2{,}47]=-3.$$

График изображён на рис. 85. Если n — целое число, то

$$f(n-0)=n-1,\quad f(n)=f(n+0)=n.$$

Пример 2. Определим функцию $f(x)$ в промежутке $0\leqslant x\leqslant 1$ — сначала для множества E дробных значений x со знаменателем вида 2^n — следующим образом. Пусть

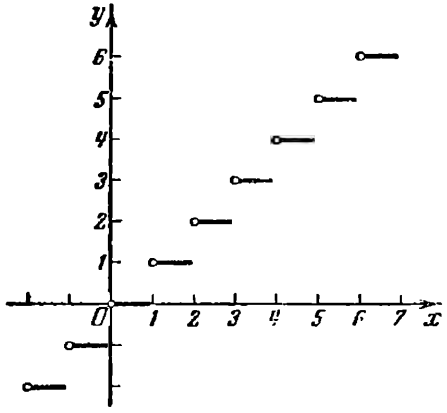


Рис. 85.

$$f(0)=0,\quad f(1)=1;$$

$$f\left(\frac{1}{2}\right)=\frac{f(0)+3f(1)}{4};$$

затем

$$f\left(\frac{1}{4}\right)=\frac{f(0)+3f\left(\frac{1}{2}\right)}{4},$$

$$f\left(\frac{3}{4}\right)=\frac{f\left(\frac{1}{2}\right)+3f(1)}{4}$$

и т. д. Вообще пусть при m нечётном

$$f\left(\frac{2m+1}{2^{n+1}}\right)=\frac{f\left(\frac{m}{2^n}\right)+3f\left(\frac{m+1}{2^n}\right)}{4}$$

(определение $f(x)$ в точках $x=a$ промежутка $0\leqslant x\leqslant 1$, не принадлежащих E, может быть дано, далее, «по непрерывности» — см. § 51).

На рис. 86, *а* построены все точки графика, соответствующие значениям x из E, для которых $n\leqslant 4$.

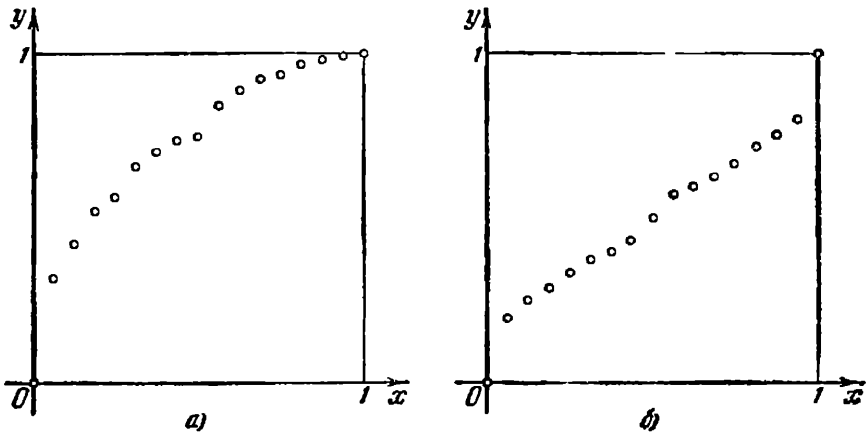


Рис. 86.

Пример 3. Приведём аналогичный пример для того же промежутка: пусть

$$f(0)=0,\quad f(1)=1;$$

затем, если вообще

$$f\left(\frac{m}{2^n}\right)=\frac{p}{q},\quad f\left(\frac{m+1}{2^n}\right)=\frac{p'}{q'},$$

то полагаем

$$f\left(\frac{2m+1}{2^{n+1}}\right)=\frac{p+p'}{q+q'}$$

$\left(\text{допускаем, что } \frac{p}{q} \text{ и } \frac{p'}{q'} \text{ — несократимые дроби, } 0=\frac{0}{1}\right)$.

Таким образом,

$$f\left(\frac{1}{2}\right)=\frac{0+1}{1+1}=\frac{1}{2},$$

$$f\left(\frac{1}{4}\right)=\frac{0+1}{1+2}=\frac{1}{3}, \qquad f\left(\frac{3}{4}\right)=\frac{1+1}{2+1}=\frac{2}{3},$$

$$f\left(\frac{1}{8}\right)=\frac{0+1}{1+3}=\frac{1}{4}, \quad f\left(\frac{3}{8}\right)=\frac{1+1}{3+2}=\frac{2}{5}, \quad f\left(\frac{5}{8}\right)=\frac{1+2}{2+3}=\frac{3}{5}, \quad f\left(\frac{7}{8}\right)=\frac{2+1}{3+1}=\frac{3}{4}$$

и т. д. (рис. 86, *б*) (здесь также возможно «продолжение на весь промежуток $0\leqslant x\leqslant 1$ по непрерывности»).

Пример 4. Функция $f(x)$ определена в промежутке $-1\leqslant x\leqslant+1$ следующими условиями:

1) $f(0)=0$;

2) $f(x)=x$, если x имеет вид $\frac{1}{2n}$ (n — целое) и $f(x)=2x$, если x имеет вид $\frac{1}{2n+1}$;

3) функция $f(x)$ в каждом из промежутков $\frac{1}{n+1}\leqslant x\leqslant\frac{1}{n}$ $(n=1,2,3,\ldots)$ — линейная;

4) функция $f(x)$ — чётная.

Читатель построит график функции $f(x)$ и убедится, что эта функция имеет минимум в точке $x=0$, хотя и нельзя указать такое число $\varepsilon(>0)$, чтобы в промежутке $-\varepsilon<x<0$ функция убывала, а в промежутке $0<x<\varepsilon$ она возрастала [1]).

Пример 5. Функции $f_1(x)$ и $f_2(x)$ при $x\neq 0$ определены равенствами

$$f_1(x)=2^{-\frac{1}{x^2}}, \quad f_2(x)=\frac{\sin x}{x}$$

и дополнительными условиями

$$f_1(0)=0, \quad f_2(0)=1.$$

И та и другая непрерывны при всех значениях x (см. § 41, пример 7 и § 43).

Пример 6. На стр. 15 мы рассматривали уравнение

$$x^2+y^2=1$$

и установили, что в промежутке $-1\leqslant x\leqslant+1$ имеются две функции $y=f(x)$, ему тождественно удовлетворяющие, именно,

$$f_1(x)=\sqrt{1-x^2} \quad \text{и} \quad f_2(x)=-\sqrt{1-x^2}.$$

---

[1]) Другие примеры, обладающие тем же свойством:

$$f(x)=2|x|+x\sin\frac{1}{x}, \quad f(x)=x^{\frac{2}{3}}+x\sin\frac{1}{x} \qquad (x\neq 0)$$

при дополнительном условии $f(0)=0$.

Вернёмся ещё раз к этому вопросу и постараемся выяснить: с к о л ь к о существует в названном промежутке функций, удовлетворяющих нашему уравнению? только ли две?

С новой точки зрения, установленной в настоящем параграфе, следует признать, что таких функций — бесконечное множество. Вот ещё одна функция, удовлетворяющая выставленному требованию:

$$f_3(x)=\begin{cases}\sqrt{1-x^2} & \text{при } -1\leqslant x\leqslant 0,\\ -\sqrt{1-x^2} & \text{при } 0<x\leqslant +1.\end{cases}$$

Эта функция — не элементарная; притом она имеет разрыв в точке $x=0$.

Представим себе, что все точки промежутка $[-1, +1]$ разбиты на два каких-либо множества A и B, и пусть функция $f(x)$ определяется условиями:

$$f(x)=\begin{cases}\sqrt{1-x^2}, & \text{если } x \text{ принадлежит множеству } A,\\ -\sqrt{1-x^2}, & \text{если } x \text{ принадлежит множеству } B.\end{cases}$$

Такая функция тоже удовлетворяет нашему уравнению. Но н е п р е р ы в н ы х функций интересующего нас класса — только две: $f_1(x)$ и $f_2(x)$.

Изложенный ход мыслей находится в соответствии с историческим процессом. В историческом плане предельный переход вошёл в математический обиход не в общем (явном) виде, а в форме суммирования рядов функций. Работы Фурье (в первой четверти прошлого столетия) создали предпосылки для расширения понятия функции: в них был указан приём представления в аналитической форме (в виде формул, содержащих предельный переход) таких функций, которые раньше считались аналитически непредставимыми и потому не подлежащими математическому исследованию. После этого введение нового определения, ныне общепринятого, стало неизбежным: оно было впервые сформулировано Лобачевским (см. стр. 12) и Леженом-Дирихле и использовано Риманом.

Числовое множество E — совокупность тех значений x, которым по определению ставятся в соответствие числовые значения y, обыкновенно называют *множеством определения* функции. Чаще всего в качестве множества определения функции приходится встречаться с *промежутками*: следует различать промежутки *замкнутые*, вида $a\leqslant x\leqslant b$ (*сегменты*), и промежутки *открытые*, вида $a<x<b$ (*интервалы*). В замкнутом промежутке $a\leqslant x\leqslant b$ имеется наибольшее число b и наименьшее a; в открытом — нет ни того, ни другого. Иногда случается иметь дело с промежутками, которые замкнуты с одного конца и открыты с другого (например, промежуток $-1<x\leqslant 1$).

Бесконечные промежутки с той стороны, на которой стоит знак бесконечности, считаются открытыми.

В настоящее время становятся общеупотребительными обозначения, которыми мы отчасти уже пользовались:

$[a, b]$ — для замкнутого промежутка $a \leqslant x \leqslant b$,

(a, b) — для открытого промежутка $a < x < b$,

$[a, b)$ — для полузамкнутого промежутка $a \leqslant x < b$ и т. п.

Примечание. Читатель, вероятно, обратил внимание на аналогию между определением числовой последовательности в § 35 и определением функции в настоящем параграфе. Однако здесь — не только аналогия, но и обобщение: задать последовательность чисел $\{a_n\}$ значит с каждым целым положительным числом n сопоставить некоторое число a_n; таким образом, *последовательность, с точки зрения настоящего определения, можно рассматривать как функцию, заданную на множестве всех целых положительных чисел.*

Как видно из приведённых примеров, «множества определения» функций могут быть весьма разнообразных типов; всё же в дальнейшем нам придётся иметь дело преимущественно с промежутками — замкнутыми или незамкнутыми.

§ 47. Свойства непрерывных функций

Чтобы отдать себе отчёт в степени общности понятия функции как «соответствия», постараемся представить себе, какова геометрическая сторона дела. С каждым значением x (хотя бы из некоторого промежутка) сопоставляется некоторое значение y: геометрически это означает, что на каждой вертикальной прямой отмечается согласно заданиому правилу одна какая-то точка. (На самом деле нельзя, конечно, отметить всех точек, которые таким образом могут быть получены; но можно «представить себе», что это сделано.) Из наличия упомянутого правила логически не вытекает, что все отмеченные точки в своей совокупности образуют «плавную кривую» (см. хотя бы пример 11 в § 45); график будет плавной, или непрерывной кривой только в случае, если функция $y = f(x)$ обладает свойством непрерывности [1]) (в рассматриваемом промежутке). Хотя в математике изучение функций, имеющих «разрывы», оказывается полезным не только в теоретическом отношении, но и с точки зрения многих приложений, тем не менее непрерывные функции (и их графики — непрерывные кривые) представляют особенно важный класс функций, обладающих замечательными свойствами. С некоторыми из этих свойств мы познакомимся теперь же; наиболее же существенное из них будет установлено в §§ 49—50.

Теорема I. *Если функция, непрерывная в некотором промежутке, в двух точках этого промежутка принимает значения разных знаков, то в какой-то точке* (хотя бы одной) *между двумя упомянутыми она принимает значение нуль.*

Эта теорема лежит в основе приближённого метода решения уравнений, посредством «проб», и доказательство следует за этим методом.

Доказательство. Пусть $f(x)$ непрерывна в промежутке I; a и b $(a < b)$ — точки этого промежутка, причём $f(a) < 0, f(b) > 0$.

[1]) См. § 42. Понятие непрерывности, очевидно, не связано с тем, задана ли функция «аналитически» или «посредством правила».

Среди всех десятичных дробей «ранга 1» (см. § 36), заключённых в промежутке $a \leqslant x \leqslant b$, можно выбрать (не обязательно—единственным способом) две рядом стоящие x_1' и x_1'' $(x_1' < x_1'')$, удовлетворяющие условиям $f(x_1') < 0$, $f(x_1'') > 0$ [1]). Затем среди всех дробей «ранга 2», заключённых в промежутке $x_1' \leqslant x \leqslant x_1''$, выберем две рядом стоящие x_2' и x_2'' $(x_2' < x_2'')$, удовлетворяющие условиям $f(x_2') < 0$, $f(x_2'') > 0$ и т. д. Последовательности $\{x_n'\}$ и $\{x_n''\}$ — монотонные и ограниченные; значит (§ 44), каждая из них имеет предел. Этот предел — один и тот же для обеих последовательностей, так как

$$x_n'' - x_n' = 10^{-n} \to 0.$$

Обозначим его через ξ. По свойству непрерывности (§ 42) из соотношений

$$x_n' \to \xi \quad \text{и} \quad x_n'' \to \xi$$

следуют соотношения

$$f(x_n') \to f(\xi) \quad \text{и} \quad f(x_n'') \to f(\xi).$$

Так как $f(x_n') < 0$ при любом n, то $f(\xi) \leqslant 0$ (§ 38, стр. 172); точно так же из $f(x_n'') > 0$ следует $f(\xi) \geqslant 0$. Итак,

$$f(\xi) = 0.$$

Следствие (обобщение). *Функция, непрерывная в некотором промежутке, не может «перейти» от одного значения к другому, не приняв любого промежуточного значения.*

Другими словами, если $f(a) = A$, $f(b) = B$, причём $a < b$ и $A \neq B$, то, каково бы ни было число X, заключённое между A и B, существует (по крайней мере одно) такое значение $x = \xi$, что 1) $a < \xi < b$, 2) $f(\xi) = X$.

Для доказательства достаточно применить доказанную теорему 1 к новой, также непрерывной, функции

$$f_1(x) \equiv f(x) - X.$$

Это свойство непрерывной функции, казалось бы, выразительнее характеризует «плавность» графика, чем само определение непрерывности. Однако оно не равносильно свойству, обычно принимаемому в качестве определения непрерывности; например, функция, заданная равенствами

$$f(x) = \begin{cases} \sin \dfrac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0, \end{cases}$$

[1]) Если бы нашлась такая дробь x_1, что $f(x_1) = 0$, то теорема была бы уже доказана: поэтому мы предполагаем, что таких дробей нет.

обладает рассматриваемым свойством и, однако, имеет точку разрыва $x=0$ (см. § 41, пример 4).

Доказанный выше результат принадлежит чешскому математику Больцано (1817 г.).

Опираясь на теорему Больцано, можно доказать существование единственного положительного действительного числа (бесконечной десятичной дроби), квадрат которого равен 2. Функция $f(x)=x^2$ непрерывна (§ 43) и возрастает в промежутке $0<x\leqslant 2$ от 0 до 4; значит, она в некоторой точке принимает значение 2:

$$f(\xi)=\xi^2=2.$$

Такое число ξ — только одно; если бы существовало два таких числа ξ' и ξ'' ($\xi'\neq\xi''$), то одно из них было бы меньше другого, например, $\xi'<\xi''$, и отсюда следовало бы $\xi'^2<\xi''^2$, а это противоречило бы равенствам $\xi'^2=2$ и $\xi''^2=2$. Итак, существует одно и только одно положительное число ξ такое, что $\xi^2=2$: оно обозначается через $\sqrt{2}$ (арифметический корень).

Точно так же доказывается существование других «не извлекающихся нацело» радикалов, затем логарифмов (исходя из непрерывности показательной функции на всей оси), обратных круговых функций, корней уравнений вроде $x^5+x-a=0$ ($a\gtrless 0$) и т. п. Таким же точно образом устанавливается, вообще говоря, существование (в смысле «соответствия») обратной функции $x=g(y)$ по отношению к любой непрерывной и монотонной функции $y=f(x)$ (см. ниже § 52).

Теорема II. *Функция, непрерывная во всех точках замкнутого промежутка, ограничена в этом промежутке.*

Доказательство. Пусть $f(x)$ непрерывна в промежутке $a\leqslant x\leqslant b$. Требуется установить, например, что существует такое число M, что при всех значениях x из этого промежутка

$$f(x)\leqslant M.$$

Допустим, что такого числа нет; тогда, каково бы ни было M, можно указать такое значение $x\equiv x_M$ в промежутке, что

$$f(x)>M.$$

Возьмём последовательность таких чисел $\{M_n\}$, что $M_n\to\infty$, и к каждому числу M_n подберём такое число x_n из промежутка, что

$$f(x_n)>M_n.$$

Тогда

$$f(x_n)\to\infty. \tag{15}$$

Из ограниченной последовательности $\{x_n\}$ выделим сходящуюся последовательность $\{x_{p_n}\}$ (см. § 38, стр. 167):

$$x_{p_n}\to\xi \qquad (a\leqslant\xi\leqslant b)^1).$$

---

[1]) В самом деле, из $a\leqslant x_{p_n}\leqslant b$ следует $a\leqslant\xi\leqslant b$ (§ 38).

По свойству непрерывности функции $f(x)$ из последнего соотношения вытекает, что последовательность $\{f(x_{p_n})\}$ стремится к конечному пределу:

$$f(x_{p_n}) \to f(\xi).$$

С другой стороны, из соотношения (15) следует, что

$$f(x_{p_n}) \to \infty.$$

Получается противоречие. Значит, число M, обладающее требуемым свойством, существует.

Заметим, что теорема неверна, если отбросить требование замкнутости промежутка. Это демонстрируется примером функции

$$f(x) = \frac{1}{x},$$

рассматриваемой в открытом (слева) промежутке $0 < x \leqslant 1$: во всех точках этого незамкнутого промежутка функция непрерывна и, однако, как легко понять, не ограничена сверху.

Теорема III (Вейерштрасса). *Среди значений, принимаемых функцией, непрерывной в замкнутом промежутке, можно указать наибольшее и наименьшее*[1]).

Другими словами: если функция $f(x)$ непрерывна в промежутке $a \leqslant x \leqslant b$, то можно указать такое число ξ_1 (не обязательно единственное), что 1) $a \leqslant \xi_1 \leqslant b$ и 2) для всех значений x из промежутка справедливо неравенство $f(x) \leqslant f(\xi_1)$; и можно указать такое число ξ_2 (тоже не обязательно единственное), что 1) $a \leqslant \xi_2 \leqslant b$ и 2) для всех значений x из промежутка $f(x) \geqslant f(\xi_2)$.

Доказательство. Установим, например, как найти число ξ_1.

По теореме II функция ограничена в рассматриваемом промежутке:

$$m < f(x) < M.$$

Среди десятичных дробей «ранга 1», очевидно, можно указать две такие, рядом стоящие y_1' и y_1'' $(y_1' < y_1'')$, что неравенство

$$f(x) < y_1''$$

выполняется для всех значений x из промежутка, а неравенство

$$f(x) < y_1'$$

— не для всех. Значит, можно указать такое значение $x = x_1$ в промежутке, что

$$y_1' \leqslant f(x_1) < y_1''.$$

---

[1]) Наибольшее из значений функции в промежутке определяется, как такое значение, которое не меньше всех прочих; наименьшее — как такое, которое не больше всех прочих. Здесь имеется аналогия с символами $\max\{a_1, \ldots, a_n\}$ и $\min\{a_1, \ldots, a_n\}$ (см. § 35).

Таким же образом среди десятичных дробей ранга 2 можно найти две такие, рядом стоящие y_2' и y_2'' $(y_2' < y_2'')$, что неравенство

$$f(x) < y_2''$$

выполняется для всех x из промежутка, а неравенство

$$f(x) < y_2',$$

— не для всех. Значит, существует такое x_2 в промежутке, что

$$y_2' \leqslant f(x_2) < y_2''$$

и т. д.

Очевидно, последовательности $\{y_n'\}$ и $\{y_n''\}$ — монотонные и ограниченные; следовательно, каждая из них имеет предел. Этот предел — один и тот же для обеих последовательностей, так как

$$y_n'' - y_n' = 10^{-n} \to 0. \qquad (16)$$

Обозначим его через η, так что

$$y_n' \to \eta \qquad (17)$$

и

$$y_n'' \to \eta. \qquad (18)$$

По нашему построению, каково бы ни было x из промежутка, и при любом целом положительном n, мы имеем

$$f(x) < y_n''.$$

Переход к пределу при $n \to \infty$ даёт

$$f(x) \leqslant \eta.$$

Нам остаётся доказать, что существует такое число $\xi_1 \equiv \xi$ в промежутке, что

$$f(\xi) = \eta.$$

По построению при любом n целом положительном существует такое x_n в промежутке, что

$$y_n' \leqslant f(x_n) < y_n'',$$

или, принимая во внимание равенство (16),

$$f(x_n) < y_n' \leqslant f(x_n) + 10^{-n}.$$

Из ограниченной последовательности $\{x_n\}$ «выделим» сходящуюся последовательность $\{x_{p_n}\}$:

$$x_{p_n} \to \xi \qquad (a \leqslant \xi \leqslant b).$$

Очевидно, имеет место неравенство

$$f(x_{p_n}) < y_{p_n}'' \leqslant f(x_{p_n}) + 10^{-p_n}.$$

Перейдём в нём к пределу (при $n \to \infty$). Принимая во внимание, что по свойству непрерывности функции $f(x)$ в точке $x = \xi$

$$f(x_{p_n}) \to f(\xi)$$

и притом вследствие (18)

$$y''_{p_n} \to \eta,$$

мы получим:

$$f(\xi) \leqslant \eta \leqslant f(\xi),$$

откуда следует:

$$f(\xi) = \eta,$$

что и требовалось доказать.

Заметим, что теорема о наибольшем значении неверна, если отбросить требование замкнутости промежутка. Примером может служить та же функция $f(x) = \frac{1}{x}$ в промежутке $0 < x \leqslant 1$, или, ещё проще, функция $f(x) = x$ в промежутке $0 \leqslant x < 1$.

Она неверна также, если отбросить требование непрерывности функции $f(x)$. Примером может служить функция, определённая равенствами

$$f(x) = \begin{cases} x & \text{при } 0 \leqslant x < 1, \\ 0 & \text{при } x = 1. \end{cases}$$

Эта функция, имеющая разрыв в точке $x = 1$, не имеет наибольшего значения в замкнутом промежутке, в котором она определена.

Теорема IV (о «равномерной» непрерывности). *Если функция $f(x)$ непрерывна в замкнутом промежутке I, то при всяком $\varepsilon\,(>0)$ можно указать такое число $\delta \equiv \delta_\varepsilon\,(>0)$, что, каковы бы ни были числа x' и x'' из I, неравенство*

$$|x' - x''| < \delta \tag{19}$$

влечёт за собой неравенство

$$|f(x') - f(x'')| < \varepsilon. \tag{20}$$

Предоставляем читателю обдумать геометрический смысл этой теоремы.

Доказательство. Пусть теорема неверна: пусть не со всяким положительным числом ε можно сопоставить число δ, обладающее указанным свойством. В таком случае существует некоторое положительное число ε^*, обладающее тем свойством, что, как бы мало ни было $\delta\,(>0)$, можно указать такие числа x' и x'' из I, что неравенство

$$|x' - x''| < \delta$$

выполняется, а неравенство

$$|f(x') - f(x'')| < \varepsilon^*$$

— не выполняется, и следовательно, имеет место противоположное неравенство

$$|f(x') - f(x'')| \geqslant \varepsilon^*.$$

Возьмём тогда последовательность положительных чисел $\{\delta_n\}$, обладающую свойством

$$\delta_n \to 0,$$

и к каждому n подберём в промежутке I пару чисел x'_n и x''_n, удовлетворяющих неравенствам

$$|x'_n - x''_n| < \delta_n \tag{21}$$

и

$$|f(x'_n) - f(x''_n)| > \varepsilon^*. \tag{22}$$

Из ограниченной последовательности $\{x'_n\}$ выделим сходящуюся последовательность $\{x'_{p_n}\}$:

$$x'_{p_n} \to \xi, \tag{23}$$

причём ξ, разумеется, принадлежит I. Из соотношений (21) и (23) следует, что также

$$x''_{p_n} \to \xi. \tag{24}$$

По свойству непрерывности соотношения (23) и (24) влекут за собой соотношения

$$f(x'_{p_n}) \to f(\xi),$$
$$f(x''_{p_n}) \to f(\xi).$$

Вследствие (22) имеет место неравенство

$$|f(x'_{p_n}) - f(x''_{p_n})| \geqslant \varepsilon^*;$$

переходя в нём к пределу при $n \to \infty$, получаем противоречие, так как левая часть стремится к нулю, тогда как правая есть постоянное положительное число.

Примечание 1. Существенное различие понятий «равномерной» непрерывности от «обыкновенной» заключается в том, что в случае «обыкновенной непрерывности в точке $x = x_0$» число δ в неравенстве $|x - x_0| < \delta$ зависит не только от числа ε в неравенстве $|f(x) - f(x_0)| < \varepsilon$, но также ещё и от значения x_0:

$$\delta \equiv \delta(\varepsilon, x_0);$$

в случае же «равномерной непрерывности в данном промежутке» число δ в неравенстве (19) зависит только от числа ε в неравенстве (20) и, конечно, ещё от самого промежутка.

Примечание 2. Из равномерной непрерывности функции в каком угодно (замкнутом или незамкнутом) промежутке I, очевидно, вытекает её непрерывность (в обычном смысле) в любой точке x_0 промежутка I.

Действительно, если неравенство (20) справедливо при каких угодно x' и x'', удовлетворяющих соотношению (19), то, в частности, можно положить $x''=x_0$ и заменить x' через x; тогда получится, что из неравенства

$$|x-x_0|<\delta$$

следует неравенство

$$|f(x)-f(x_0)|<\varepsilon,$$

как полагается по обычному определению.

Примечание 3. Теорема IV неверна в случае незамкнутого промежутка. Об этом свидетельствует всё тот же пример: $f(x)=\frac{1}{x}$ $(0<x\leqslant 1)$. Полагая здесь $\varepsilon=1$, $x'=\delta$, $x''=\frac{\delta}{2}$, мы получаем при $\delta<1$

$$|x'-x''|=\frac{\delta}{2}<\delta$$

и, однако,

$$|f(x')-f(x'')|=\left|\frac{1}{\delta}-\frac{2}{\delta}\right|=\frac{1}{\delta}>1.$$

Таким образом, *в случае замкнутого промежутка свойства обыкновенной непрерывности в промежутке и равномерной непрерывности в промежутке строго эквивалентны;* в случае незамкнутого промежутка это не так.

§ 48. Равномерная сходимость последовательности непрерывных функций

В § 45 понятие предела последовательности функции было сведено к понятию последовательности чисел: требовалась сходимость последовательности $\{f_n(x)\}$ в отдельности для каждого значения x из рассматриваемого промежутка I.

Можно подойти к тому же вопросу иначе, рассматривая функцию в промежутке I как целое. Ниже будет изложен один из таких возможных подходов. При этом мы условимся, что на первых порах будем иметь дело лишь с непрерывными функциями, рассматриваемыми в одном и том же замкнутом промежутке [1]).

Рассмотрим две какие-нибудь непрерывные функции $f(x)$ и $g(x)$ в замкнутом (и следовательно, конечном) промежутке $I(a\leqslant x\leqslant b)$. Мы укажем способ измерять «отклонение» одной функции от другой, или одного графика от другого, посредством некоторой числовой характеристики. Когда идёт речь об «отклонении» одного числа a от другого b, то под таковым мы понимаем (см. § 36) абсолютную величину их разности $|a-b|$. В качестве же «меры отклонения» (или просто «отклонения») функции $f(x)$ от функции $g(x)$ в про-

[1]) Такое ограничение лишь отчасти связано с существом дела; главным же образом оно оправдывается нашим намерением сосредоточить внимание на простом и особенно важном случае.

межутке I мы примем *максимум* (наибольшее значение) *абсолютной величины их разности в этом промежутке:*

$$\rho(f, g) \equiv \rho_I(f, g) = \max_I |f(x) - g(x)|. \quad (25)$$

Геометрически, как легко понять, отклонение $\rho(f, g)$ представляет собой наибольшую из длин вертикальных отрезков, соединяющих точки графиков f и g с одинаковыми абсциссами (рис. 87). Так как функции $f(x)$ и $g(x)$ непрерывны, то непрерывна также их разность $f(x) - g(x)$, а также и её абсолютная величина $|f(x) - g(x)|$ [1]); следовательно (по теореме III предыдущего параграфа), среди упомянутых отрезков есть наибольший.

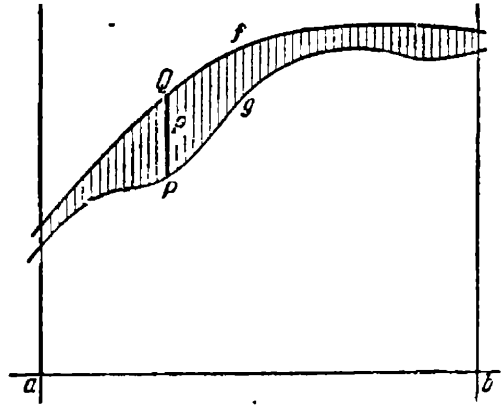


Рис. 87.

Итак, отклонение $\rho(f, g)$ обладает следующим свойством: каково бы ни было значение x из промежутка I, справедливо соотношение

$$|f(x) - g(x)| \leqslant \rho(f, g), \quad (26)$$

причём равенство имеет место хотя бы для одного из этих значений.

Отклонение $\rho(f, g)$ всегда неотрицательно:

$$\rho(f, g) \geqslant 0 \quad (27)$$

и равно нулю в том и только в том случае, если $f(x) \equiv g(x)$ в промежутке I (графики совпадают).

Обратимся теперь к определению равномерной сходимости последовательности функции.

Говорят, что *последовательность функций*

$$\{f_n(x)\} \equiv f_1(x), \ f_2(x), \ \ldots, \ f_n(x), \ \ldots,$$

непрерывных в одном и том же замкнутом промежутке I, в этом промежутке сходится (стремится) равномерно к непрерывной же функции $f(x)$, если предел числовой последовательности

$$\{\rho(f_n, f)\} \equiv \rho(f_1, f), \ \rho(f_2, f), \ \ldots, \ \rho(f_n, f), \ \ldots$$

равен нулю:

$$\rho(f_n, f) \to 0. \quad (28)$$

[1]) Мы предоставляем читателю в данной связи доказать теорему:
Если непрерывна в точке $x = c$ функция $F(x)$, то непрерывна также и её абсолютная величина $|F(x)|$.

Для обозначения равномерной сходимости мы будем иногда пользоваться записью:

$$f_n \rightrightarrows f. \tag{29}$$

Теорема. *Из равномерной сходимости* ($f_n \rightrightarrows f$) *следует сходимость в обычном* (см. § 45) *смысле* ($f_n \to f$).

В самом деле, соотношение (29), равносильное (28), означает, что, как бы мало ни было $\varepsilon\,(>0)$, при достаточно больших значениях n справедливо неравенство

$$\rho(f_n, f) < \varepsilon. \tag{30}$$

Пусть x — какое-нибудь число из промежутка I. По свойству (26)

$$|f_n(x) - f(x)| \leqslant \rho(f_n, f); \tag{31}$$

значит,

$$|f_n(x) - f(x)| < \varepsilon.$$

Итак, при любом x из I

$$f_n(x) \to f(x).$$

(Доказательство можно резюмировать следующим образом: раз наибольший из «отрезков» $|f_n(x) - f(x)|$ стремится к нулю при $n \to \infty$, то и каждый из них в отдельности также стремится к нулю.)

Приведём примеры, показывающие, что, обратно, из обыкновенной сходимости не следует равномерная сходимость.

Пример 1.

$$f_n(x) = \frac{1}{1 + n^2\left(x - \frac{1}{\sqrt{n}}\right)^2} \quad (n = 1, 2, 3, \ldots). \tag{32}$$

Значения x здесь произвольны; ограничимся хотя бы промежутком

$$0 \leqslant x \leqslant 1.$$

Мы получаем:

1) при $x = 0$

$$f_n(0) = \frac{1}{1+n} \to 0,$$

2) при $x \neq 0$

$$f_n(x) = \frac{1}{n^2x^2} \cdot \frac{1}{\left(1 - \frac{1}{x\sqrt{n}}\right)^2 + \frac{1}{n^2x^2}} \to 0.$$

Таким образом, при всех значениях x

$$f_n(x) \to 0,$$

так что предельная функция $f(x)$ тождественно равна нулю. И тем не менее равномерной сходимости нет, так как

$$\rho(f_n, f) = \max_I |f_n(x) - f(x)| = \max_I f_n(x) = 1\ ^1).$$

---

[1]) Непосредственно из формулы видно, что $f_n(x)$ принимает наибольшее значение при $x = \frac{1}{\sqrt{n}}$.

Полезно отдать себе отчёт в геометрическом смысле этого примера: график функции $f_n(x)$ тот же, что и в примере 3, § 45 (см. рис. 80), но сдвинут вправо на расстояние $\frac{1}{\sqrt{n}}$.

В следующем примере функции последовательности не принадлежат к числу элементарных, но геометрическая сторона дела более проста.

Пример 2. В промежутке $0 \leqslant x \leqslant 1$ непрерывные функции $f_n(x)$ заданы условиями:

$$f_n(x) = \begin{cases} nx & \text{при } 0 \leqslant x \leqslant \frac{1}{n}, \\ 2 - nx & \text{при } \frac{1}{n} \leqslant x \leqslant \frac{2}{n} \\ 0 & \text{при } \frac{2}{n} \leqslant x \leqslant 1 \end{cases} \quad (n = 1, 2, 3 \ldots), \tag{33}$$

(на рис. 88 изображён график f_{10}). Мы получаем:

1) при $x = 0$

$$f_n(0) = 0 \to 0,$$

2) при $0 < x \leqslant 1$:

$f_n(x) = 0$, если только $n \geqslant \frac{2}{x}$;

значит, и в этом случае

$$f_n(x) \to 0.$$

Итак, в нашем промежутке

$$f_n(x) \to f(x) \equiv 0.$$

Тем не менее,

$$\rho(f_n, f) = \max f_n(x) = f_n\left(\frac{1}{n}\right) = 1.$$

Поэтому сходимость — не равномерная.

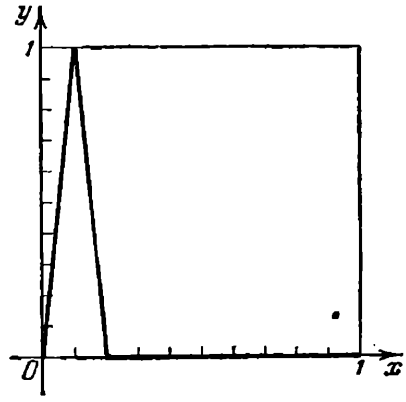


Рис. 88.

Геометрический смысл равномерной сходимости чрезвычайно прост. Соотношение (29) означает, что, как бы мало ни было ε, при достаточно больших значениях n имеет место неравенство

$$\rho(f_n, f) < \varepsilon;$$

это же неравенство по отмеченному выше свойству отклонения равносильно одновременному выполнению бесконечного множества неравенств

$$|f_n(x) - f(x)| < \varepsilon \tag{34}$$

(где x может принимать любое значение из промежутка I).

Придавая соотношению (34) вид

$$f(x) - \varepsilon < f_n(x) < f(x) + \varepsilon,$$

мы можем следующим образом истолковать равномерную сходимость

$$f_n \rightrightarrows f.$$

Вообразим графики функций $f(x)-\varepsilon$ и $f(x)+\varepsilon$, получающиеся из графика $f(x)$ посредством перенесения графика вверх и вниз на отрезок ε. Тогда при достаточно больших значениях $n\,(n>n_\varepsilon)$ график функции $f_n(x)$ весь заключён в «полосе» между этими двумя графиками (рис. 89).

Если бы мы условились эту полосу называть «ε-окрестностью» графика $f(x)$, то можно было бы сказать иначе, что графики «почти всех» функций $f_n(x)$ попадают целиком в эту «ε-окрестность», т. е. лишь конечное число их не попадает в неё (см. определение предела числовой последовательности, приведённое на стр. 164).

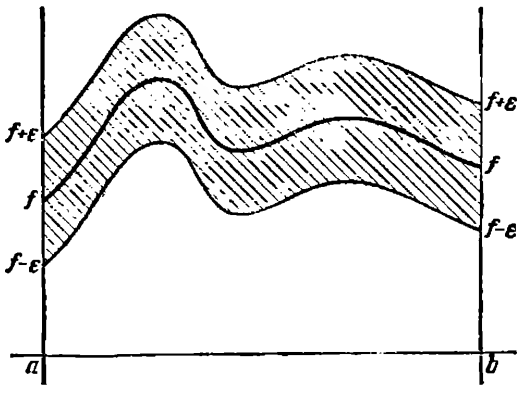


Рис. 89.

Выше равномерная сходимость $f_n(x)\rightrightarrows f(x)$ была определена в предположении, что все функции последовательности $\{f_n\}$, а также и предельная функция $f(x)$ в рассматриваемом промежутке I непрерывны. В таком ограничении нет логической необходимости. В самом деле, из приведённого определения равномерной сходимости вытекает: *как бы мало ни было* $\varepsilon\,(>0)$, *существует такое* n (*зависящее только от* ε, *но не от* x), *что при всяком значении* x *из промежутка* I *справедливо неравенство* (34). Эту словесную формулу обыкновенно и принимают в качестве определения равномерной сходимости, не требуя в самом определении непрерывности функций $f_n(x)$ и $f(x)$.

При таком понимании равномерной сходимости сформулируем теорему:

Если последовательность функций $\{f_n(x)\}$, *заданных и непрерывных в промежутке* I, *сходится в нём равномерно к некоторой функции* $f(x)$, *то эта функция также непрерывна в* I.

Доказательство. Пусть c — какое угодно значение переменной из промежутка I, ε — заданное сколь угодно малое положительное число. Из тождества

$$f(x)-f(c)=[f(x)-f_n(x)]+[f_n(x)-f_n(c)]+[f_n(c)-f(c)]$$

следует неравенство

$$|f(x)-f(c)|\leqslant$$
$$\leqslant|f(x)-f_n(x)|+|f_n(x)-f_n(c)|+|f_n(c)-f(c)|. \qquad (35)$$

Выберем в нём: 1) n настолько большим, чтобы при каких угодно x из I (и, в частности, при $x=c$) имело место неравенство

$$|f(x)-f_n(x)|<\frac{\varepsilon}{3},$$

и затем 2) $\delta(>0)$ настолько малым, чтобы при выбранном значении n и при условии $|x-c|<\delta$ осуществлялось неравенство

$$|f_n(x)-f_n(c)|<\frac{\varepsilon}{3};$$

тогда каждое из трёх слагаемых в правой части неравенства (35) будет меньше чем $\frac{\varepsilon}{3}$ и, следовательно, при условии $|x-c|<\delta$ мы будем иметь

$$|f(x)-f(c)|<\varepsilon.$$

Отсюда видно (см. § 42, стр. 189), что

$$\lim_{x\to c} f(x)=f(c),$$

т. е. что функция $f(x)$ непрерывна в точке c промежутка I, а следовательно, и во всём промежутке I.

§ 49. Теорема Вейерштрасса-Бернштейна о приближении непрерывной функции с помощью рациональных многочленов

Основная задача, которой посвящены главы I—II настоящей статьи и к разрешению которой в одинаковой степени привлекается внимание школьника, студента и преподавателя, заключается в следующем: *дано уравнение, выражающее одну переменную величину y через другую x; требуется построить соответствующий график.* В этой задаче нет никаких принципиальных трудностей, дело осложнено лишь разнообразием возможных случаев и необходимостью выработать надлежащие технические приёмы.

Но и практика и исследователя часто интересует обратная задача, представляющая трудности принципиального порядка. Вкратце её можно сформулировать так: *дан график, требуется построить соответствующее уравнение.* Наблюдатель или экспериментатор, регистрирующий ход некоторого процесса, в котором каждое «состояние» характеризуется значениями двух функционально-связанных величин (переменных, или параметров), получает на координатной («изобразительной») плоскости запись процесса в виде некоторой эмпирической — или (как говорили ещё во время Эйлера) «произвольной» — кривой; чтобы приступить к математическому исследова-

нию изучаемого явления, нужно — точно или хотя бы в каком-то смысле приближённо — написать уравнение этой кривой.

Как подобрать такое уравнение?

Постановка вопроса требует уточнения, и это можно сделать различными способами.

Допустим, что заданная кривая в пределах рассматриваемого промежутка $I(a \leqslant x \leqslant b)$ имеет одну общую точку со всякой вертикальной прямой, что позволяет рассматривать ординату y каждой точки кривой, как функцию абсциссы x

$$y = f(x)$$

(функция здесь понимается, конечно, как «соответствие»); предположим дальше, что функция $f(x)$ непрерывна в точном математическом смысле [1]) (см. § 42).

Иногда выбирают на данной кривой конечное число точек

$$M_i(x_i, y_i) \quad (i = 1, 2, \ldots, m)$$

и затем так подбирают элементарную непрерывную функцию, чаще всего — рациональный многочлен $P(x)$, чтобы график в точности проходил через выбранные точки (не заботясь при этом о точном совпадении графиков в промежуточных точках). Задача нахождения этого рода функции носит название *интерполяционной*.

Другая постановка вопроса — и именно она нас в данном случае интересует — заключается в том, чтобы отклонение подбираемой элементарной функции (могочлена $P(x)$) от данной функции во всех точках промежутка было меньше заранее назначенного числа ε. Иными словами, требуют, чтобы во всём промежутке выполнялось неравенство

$$|P(x) - f(x)| < \varepsilon,$$

или (ещё иначе), в обозначениях § 48,

$$\rho(P, f) < \varepsilon.$$

Такая задача носит название *аппроксимационной* (задача *приближения* функции).

Естественно интересоваться, можно ли подобрать многочлен $P(x)$ к заданной функции $f(x)$ и заданному наперёд сколь угодно малому числу ε.

Утвердительный ответ даётся теоремой Вейерштрасса (1885 г.), которая является обратной по отношению к теореме, приведённой в конце § 48. Вот её точная формулировка:

---

[1]) В последнем требовании есть элементы идеализации: по отношению к эмпирическим кривым можно лишь условно говорить о его выполнении. Но оно выполняется в точном смысле, например, в случае «склеенных» непрерывных графиков: см. хотя бы § 45, примеры 7 или 9, при условии, что функции $u(x)$ и $v(x)$ — непрерывные, и притом $u(0) = v(0)$.

Какова бы ни была функция $f(x)$, непрерывная в замкнутом [1]) *промежутке $I(a \leqslant x \leqslant b)$, можно указать такую последовательность многочленов $\{P_n(x)\}$, которая в промежутке I равномерно сходится к $f(x)$.*

Итак, проблема представления данного «графика» посредством «уравнения» имеет точное решение

$$f(x) = \lim_{n \to \infty} P_n(x)$$

и приближённое решение

$$f(x) \sim P_n(x),$$

причём возможная погрешность $|P_n(x) - f(x)|$ не превышает числа $\rho(P_n, f) = \max |P_n(x) - f(x)|$, не зависящего от x и стремящегося к нулю при $n \to \infty$.

Остриё всякого пишущего инструмента (карандаша, пера, куска мела) — не точка, и «реальный» график — след, остающийся на бумаге, при движении пишущего инструмента — не «идеальная математическая кривая», а «полоса», имеющая некоторую «ширину». Поэтому, несколько упрощая, можно сформулировать теорему Вейерштрасса и таким образом:

Как бы «извилист» ни был данный «реальный» график, проведённый «одним движением» пишущего инструмента (см. рис. 89), *всегда можно найти рациональный многочлен, график которого «совпадает» с данным графиком.*

Возвращаясь к точной формулировке, нужно заметить, что последовательность многочленов $\{P_n(x)\}$ определяется не однозначно. Различные доказательства теоремы, предложенные различными авторами, приводят к тому или иному приёму построения приближающей последовательности.

Первоначально данные доказательства не были общедоступными. Мы приведём ниже, в свободном изложении, доказательство советского учёного, академика С. Н. Бернштейна, предложенное им в 1912 году. Будучи вполне элементарным, оно потребует, впрочем, некоторой предварительной подготовки.

Предположим, для простоты, что речь идёт о промежутке

$$0 \leqslant x \leqslant 1.$$

Многочлены Бернштейна последовательно возрастающих степеней обозначаются через $B_n(x)$. Они имеют вид:

$$B_n(x) = \sum_{m=0}^{n} f\left(\frac{m}{n}\right) C_n^m x^m (1-x)^{n-m}, \tag{36}$$

где через C_n^m обозначены биномиальные коэффициенты

$$C_n^m = \frac{n!}{m!(n-m)!};$$

с условием полагать $0! = 1$.

---

[1]) Существенное предположение.

Теорема Бернштейна. *Пусть $f(x)$ — функция, непрерывная в замкнутом промежутке $0 \leqslant x \leqslant 1$. Тогда последовательность многочленов $\{B_n(x)\}$ в этом промежутке равномерно стремится к $f(x)$:*

$$B_n(x) \rightrightarrows f(x). \tag{37}$$

Лемма. *Если $p+q=1$, то сумма*

$$S^{(n)} = \sum_{m=0}^{n} (m-np)^2 C_n^m p^m q^{n-m} \qquad (n=0,\ 1,\ 2,\ \ldots)$$

равна npq.

Доказательство. Вычислим предварительно суммы

$$S_0^{(n)} = \sum_{m=0}^{n} C_n^m p^m q^{n-m},$$

$$S_1^{(n)} = \sum_{m=0}^{n} m C_n^m p^m q^{n-m}$$

и

$$S_2^{(n)} = \sum_{m=0}^{n} m(m-1) C_n^m p^m q^{n-m}.$$

Что касается первой из них, то она представляет собой не что иное, как разложение по формуле бинома

$$S_0^{(n)} = (p+q)^n,$$

и потому

$$S_0^{(n)} = 1 \qquad (n=0,\ 1,\ 2,\ \ldots).$$

Чтобы вычислить сумму $S_1^{(n)}$, заметим предварительно, что

$$mC_n^m = nC_{n-1}^{m-1} \qquad (m \geqslant 1, \quad n \geqslant 1);$$

отсюда следует при $n \geqslant 1$:

$$S_1^{(n)} = \sum_{m=0}^{n} m C_n^m p^m q^{n-m} = \sum_{m=1}^{n} m C_n^m p^m q^{n-m} =$$

$$= n \sum_{m=1}^{n} C_{n-1}^{m-1} p^m q^{n-m} =$$

$$= np \sum_{m=1}^{n} C_{n-1}^{m-1} p^{m-1} q^{(n-1)-(m-1)} =$$

$$= np \sum_{m=0}^{n-1} C_{n-1}^{m} p^m q^{(n-1)-m} = np \cdot S_0^{(n-1)} = np.$$

Чтобы вычислить сумму S_2^n, заметим, что

$$m(m-1)\,C_n^m = n(n-1)\,C_{n-2}^{m-2} \qquad (m \geqslant 2, \quad n \geqslant 2),$$

и потому (при $n \geqslant 2$)

$$S_2^{(n)} = \sum_{m=0}^{n} m(m-1)\,C_n^m p^m q^{n-m} = \sum_{m=2}^{n} m(m-1)\,C_n^m p^m q^{n-m} =$$

$$= n(n-1) \sum_{m=2}^{n} C_{n-2}^{m-2} p^m q^{n-m} =$$

$$= n(n-1)\,p^2 \sum_{m=2}^{n} C_{n-2}^{m-2} p^{m-2} q^{(n-2)-(m-2)} =$$

$$= n(n-1)\,p^2 \sum_{m=0}^{n-2} C_{n-2}^{m} p^m q^{(n-2)-m} =$$

$$= n(n-1)\,p^2 \cdot S_0^{(n-2)} = n(n-1)\,p^2.$$

Итак, мы получаем:

$$S_0^{(n)} = 1, \quad S_1^{(n)} = np, \quad S_2^{(n)} = n(n-1)\,p^2 \qquad (n = 0,\ 1,\ 2,\ \ldots)\ ^1).$$

Переходя к вычислению суммы $S^{(n)}$, обратим внимание на тождество

$$(m-np)^2 = n^2p^2 - (2np-1)\,m + m(m-1);$$

из него вытекает (при $n \geqslant 0$) с помощью уже полученных формул:

$$S^{(n)} = \sum_{m=0}^{n} (m-np)^2\,C_n^m p^m q^{n-m} =$$

$$= \sum_{m=0}^{n} \{\, n^2p^2 - (2np-1)\,m + m(m-1) \,\}\, C_n^m p^m q^{n-m} =$$

$$= n^2p^2 \sum_{m=0}^{n} C_n^m p^m q^{n-m} - (2np-1) \sum_{m=0}^{n} m C_n^m p^m q^{n-m} +$$

$$+ \sum_{m=0}^{n} m(m-1) C_n^m p^m q^{n-m} =$$

$$= n^2p^2 \cdot S_0^{(n)} - (2np-1)\,S_1^{(n)} + S_2^{(n)} =$$

$$= n^2p^2 \cdot 1 - (2np-1)\,np + n(n-1)\,p^2 =$$

$$= np(1-p) = npq.$$

---

[1]) Справедливость этих формул для сумм $S_1^{(0)}$, $S_2^{(0)}$ и $S_2^{(1)}$ проверяется непосредственно.

§ 50. Доказательство теоремы

Условимся в сокращённом обозначении:

$$u_m^{(n)}(x) = C_n^m x^m (1-x)^{n-m} \qquad (m = 0,\ 1,\ \ldots,\ n;\ n = 0,\ 1,\ 2,\ \ldots);$$

в таком случае

$$B_n(x) = \sum_{m=0}^{n} f\left(\frac{m}{n}\right) u_m^{(n)}(x). \tag{38}$$

Полагая в формуле

$$S_0^{(n)} = \sum_{m=0}^{n} C_n^m p^m q^{n-m} = 1$$

p равным x, q равным $1-x$ (так что $p+q=1$), мы получаем тождество (относительно x):

$$\sum_{m=0}^{n} u_m^{(n)}(x) = 1, \tag{39}$$

откуда следует, что

$$f(x) = \sum_{m=0}^{n} f(x)\, u_m^{(n)}(x). \tag{40}$$

Вычитая (40) из (38), мы будем иметь:

$$B_n(x) - f(x) = \sum_{m=0}^{n} \left[f\left(\frac{m}{n}\right) - f(x)\right] u_m^{(n)}(x),$$

так что (по свойствам абсолютной величины)

$$|B_n(x) - f(x)| \leqslant \sum_{m=0}^{n} \left|f\left(\frac{m}{n}\right) - f(x)\right| u_m^{(n)}(x)\ ^{1}). \tag{41}$$

Нам нужно убедиться, что сумма в правой части при достаточно больших n становится меньше любого наперёд заданного числа $\varepsilon\ (>0)$. Для этого придётся воспользоваться свойством непрерывности функции $f(x)$.

Так как функция $f(x)$ предполагается непрерывной в замкнутом промежутке $0 \leqslant x \leqslant 1$, то она в этом промежутке 1) ограничена (см. § 47, теорема II) и 2) равномерно непрерывна (см. § 47, теорема IV). Это значит:

1° можно указать такое число M, что

$$|f(x)| \leqslant M \qquad (0 \leqslant x \leqslant 1), \tag{42}$$

---

[1]) Очевидно, $u_m^{(n)}(x) \geqslant 0$ при $0 \leqslant x \leqslant 1$.

2° к заданному числу ε можно подобрать такое число $\delta(>0)$, что, каковы бы ни были числа x' и x'' из промежутка $0 \leqslant x \leqslant 1$, из неравенства

$$|x' - x''| < \delta$$

непременно следует неравенство

$$|f(x') - f(x'')| < \frac{\varepsilon}{2}. \tag{43}$$

Перейдём теперь к «оценке» суммы, стоящей в правой части (41). Эту сумму разобьём на две суммы, и именно таким образом. При суммировании индекс m пробегает все значения от 0 до n:

$$m = 0,\ 1,\ 2,\ \ldots,\ n;$$

из этих значений одни удовлетворяют неравенству

$$\left|\frac{m}{n} - x\right| < \delta, \tag{44}$$

другие — не удовлетворяют. Разобьём сумму на две суммы: к первой отнесём те члены, для которых неравенство (44) выполняется, ко второй — те, для которых это неравенство не выполняется, и получающиеся таким образом частные суммы снабдим значками I и II [1]). Итак,

$$\sum_{m=0}^{n} \left|f\left(\frac{m}{n}\right) - f(x)\right| u_m^{(n)}(x) =$$

$$= \sum\nolimits_{\mathrm{I}} \left|f\left(\frac{m}{n}\right) - f(x)\right| u_m^{(n)}(x) + \sum\nolimits_{\mathrm{II}} \left|f\left(\frac{m}{n}\right) - f(x)\right| u_m^{(n)}(x). \tag{45}$$

Рассмотрим теперь каждую из двух сумм в отдельности.

В первой сумме значок m принимает лишь такие значения, которые удовлетворяют неравенству (44) и потому согласно пункту 2° для всех членов этой суммы справедливо неравенство

$$\left|f\left(\frac{m}{n}\right) - f(x)\right| < \frac{\varepsilon}{2}.$$

Итак, мы получаем:

$$\sum\nolimits_{\mathrm{I}} \left|f\left(\frac{m}{n}\right) - f(x)\right| u_m^{(n)}(x) \leqslant \sum\nolimits_{\mathrm{I}} \frac{\varepsilon}{2} u_m^{(n)}(x) = \frac{\varepsilon}{2} \sum\nolimits_{\mathrm{I}} u_m^{(n)}(x). \tag{46}$$

---

[1]) Нам нисколько не мешает то обстоятельство, что разбивка членов суммы $\sum_{m=0}^{n}$ на две суммы $\sum_{\mathrm{I}}$ и $\sum_{\mathrm{II}}$ зависит от числового значения x.

Обратимся на минуту к сумме $\sum_{m=0}^{n} u_m^{(n)}(x)$: её можно было бы разбить на две, точно так же, как и предыдущую,

$$\sum_{m=0}^{n} u_m^{(n)}(x) = \sum_{\text{I}} u_m^{(n)}(x) + \sum_{\text{II}} u_m^{(n)}(x);$$

принимая во внимание, что среди выражений $u_m^{(n)}(x)$ нет отрицательных и что сумма слева согласно (39) равна единице, мы заключаем, что

$$\sum_{\text{I}} u_m^{(n)}(x) = 1 - \sum_{\text{II}} u_m^{(n)}(x) \leqslant 1. \tag{47}$$

Возвращаясь к неравенству (46), мы теперь можем утверждать, что

$$\sum_{\text{I}} \left| f\left(\frac{m}{n}\right) - f(x) \right| u_m^{(n)}(x) \leqslant \frac{\varepsilon}{2}. \tag{48}$$

Перейдём теперь к рассмотрению второй суммы в правой части (45):

$$\sum_{\text{II}} \left| f\left(\frac{m}{n}\right) - f(x) \right| u_m^{(n)}(x).$$

Так как по свойствам абсолютной величины и с помощью (42)

$$\left| f\left(\frac{m}{n}\right) - f(x) \right| \leqslant \left| f\left(\frac{m}{n}\right) \right| + \left| f(x) \right| \leqslant M + M = 2M,$$

то

$$\sum_{\text{II}} \left| f\left(\frac{m}{n}\right) - f(x) \right| u_m^{(n)}(x) \leqslant \sum_{\text{II}} 2M u_m^{(n)}(x) = 2M \sum_{\text{II}} u_m^{(n)}(x). \tag{49}$$

Согласно нашему условию во всех членах суммы, отмеченной значком II, неравенство (44) не выполняется, и следовательно, выполняется противоположное неравенство

$$\left| \frac{m}{n} - x \right| \geqslant \delta,$$

которому можно придать также вид

$$|m - nx| \geqslant \delta n,$$

или

$$(m - nx)^2 \geqslant \delta^2 n^2,$$

или, наконец,

$$1 \leqslant \left(\frac{m - nx}{\delta n}\right)^2.$$

Умножим обе части последнего неравенства на неотрицательную величину $u_m^{(n)}(x)$ и просуммируем (разумеется, со значком II):

$$\sum_{\text{II}} u_m^{(n)}(x) \leqslant \sum_{\text{II}} \left(\frac{m - nx}{\delta n}\right)^2 u_m^{(n)}(x) = \frac{1}{\delta^2 n^2} \sum_{\text{II}} (m - nx)^2 u_m^{(n)} x. \tag{50}$$

Обращаясь к сумме $\sum_{m=0}^{n}(m-nx)^2 u_m^{(n)}(x)$, мы можем её разбить по тому же правилу, что и прежние суммы:

$$\sum_{m=0}^{n}(m-nx)^2 u_m^{(n)}(x)=\sum_{\text{I}}(m-nx)^2 u_m^{(n)}(x)+\sum_{\text{II}}(m-nx)^2 u_m^{(n)}(x),$$

откуда сейчас же следует неравенство

$$\sum_{\text{II}}(m-nx)^2 u_m^{(n)}(x)\leqslant\sum_{m=0}^{n}(m-nx)^2 u_m^{(n)}(x). \tag{51}$$

Сопоставляя неравенства (50) и (51), мы получаем дальше:

$$\sum_{\text{II}}u_m^{(n)}(x)\leqslant\frac{1}{\delta^2 n^2}\sum_{m=0}^{n}(m-nx)^2 u_m^{(n)}(x). \tag{52}$$

Сумму, стоящую справа, нам очень легко вычислить: для этого достаточно в формуле леммы положить $p=x$, $q=1-x$; тогда оказывается, что

$$\sum_{m=0}^{n}(m-nx)^2 u_m^{(n)}(x)=\sum_{m=0}^{n}(m-nx)^2 C_n^m x^m(1-x)^{n-m}=nx(1-x).$$

Поэтому неравенство (52) принимает вид

$$\sum_{\text{II}}u_m^{(n)}(x)\leqslant\frac{x(1-x)}{\delta^2 n}.$$

По поводу выражения $x(1-x)$ можно заметить (см. § 9), что

$$x(1-x)=\frac{1}{4}-\left(x-\frac{1}{2}\right)^2\leqslant\frac{1}{4};$$

значит,

$$\sum_{\text{II}}u_m^{(n)}(x)\leqslant\frac{1}{4\delta^2 n}.$$

Сопоставляя с этим неравенством раньше полученное неравенство (49), можно написать:

$$\sum_{\text{II}}\left|f\left(\frac{m}{n}\right)-f(x)\right|u_m^{(n)}(x)\leqslant\frac{M}{2\delta^2 n}. \tag{53}$$

Наконец, собирая вместе соотношения (41), (45) и (48), (53), мы заключаем:

$$|B_n(x)-f(x)|<\frac{\varepsilon}{2}+\frac{M}{2\delta^2 n}.$$

Число ε задано наперёд. Число δ подобрано нами в зависимости от заданного ε, принимая во внимание непрерывность функции $f(x)$. Но ничего не сказано пока о степени n многочлена $B_n(x)$. Мы допустим теперь, что степень n настолько велика, что удовлетворяет неравенству

$$\frac{M}{2\delta^2 n} < \frac{\varepsilon}{2},$$

т. е. допустим, что

$$n > \frac{M}{\delta^2 \varepsilon}. \tag{54}$$

В таком случае

$$|B_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Это неравенство имеет место при любом значении x из промежутка $0 \leqslant x \leqslant 1$. Придавая x значение, при котором непрерывная функция $|B_n(x) - f(x)|$ в этом промежутке достигает максимума (см. теорему III, § 47), мы приходим к неравенству

$$\rho(B_n, f) = \max |B_n(x) - f(x)| < \varepsilon.$$

Таким образом, ко всякому ε можно подобрать такое n_ε $\left(\text{например, } n_\varepsilon = \frac{M}{\delta^2 \varepsilon}\right)$, что при $n > n_\varepsilon$ имеет место неравенство

$$\rho(B_n, f) < \varepsilon.$$

Но это как раз означает, что

$$\rho(B_n, f) \to 0$$

или, в промежутке $0 \leqslant x \leqslant 1$,

$$B_n(x) \rightrightarrows f(x).$$

Теорема доказана.

Пример 1. Полагая $f(x) = a^x$, мы получаем:

$$B_n(x) = \sum_{m=0}^{n} a^{\frac{m}{n}} C_n^m x^m (1-x)^{n-m} = \left[(1-x) + a^{\frac{1}{n}} x\right]^n =$$

$$= \left[1 + (\sqrt[n]{a} - 1)\, x\right]^n.$$

Пример 2. Полагая $f(x) = \sin \pi x$, мы получаем [1]:

$$B_n(x) = \sum_{m=0}^{n} \sin \frac{m\pi}{n} \cdot C_n^m x^m (1-x)^{n-m} =$$

$$= \frac{1}{2i}\left\{\left[\left(1 - 2x \sin^2 \frac{\pi}{2n}\right) + ix \sin \frac{\pi}{n}\right]^n - \left[\left(1 - 2x \sin^2 \frac{\pi}{2n}\right) - ix \sin \frac{\pi}{n}\right]^n\right\}.$$

[1]) Следует принять во внимание, что

$$\sin t = \frac{e^{it} - e^{-it}}{2i}$$

(см. стр. 504). Все коэффициенты многочлена $B_n(x)$, конечно, — действительные.

Примечание. Неравенство (54) позволяет судить о том, какую степень n достаточно выбрать, чтобы сделать отклонение $\rho(B_n, f)$ меньшим, чем заданное число ε.

Например, если $f(x)=\sin \pi x$, то можно положить $M=1$, притом [1])

$$|\sin \pi x' - \sin \pi x''| < \pi |x' - x''|,$$

так что достаточно положить

$$\delta = \frac{\varepsilon}{\pi},$$

чтобы при $|x' - x''| < \delta$ иметь $|f(x') - f(x'')| < \varepsilon$.

В таком случае неравенство (54) принимает вид

$$n > \frac{\pi}{\varepsilon^3}.$$

Итак, если, например, $\varepsilon = 0{,}1$, то достаточно взять многочлен $B_n(x)$ степени большей, чем 3141, чтобы отклонение $\rho(B_n, f)$ стало меньше 0,1.

Не следует особенно огорчаться этим результатом: взять столь высокую степень достаточно для того, чтобы отклонение стало меньше 0,1, но необходимости в том нет, и требуемое приближение на самом деле достигается при гораздо меньших значениях n.

Если бы мы поставили своей задачей не выяснение принципиальной стороны дела, а оценку фактической погрешности, то рассуждение пришлось бы строить иначе (оно было бы значительно сложнее).

§ 51. Определение показательной функции. Продолжение непрерывной функции за пределы всюду плотного множества

Вопрос, который мы здесь перед собой поставим, заключается в следующем: можно ли установить значение функции $y=f(x)$ в некоторой точке $x=c$, если известны её значения в каких-то других точках, отличных от от точки c?

Ответ, конечно, должен быть категорически отрицательным, если функция $f(x)$ не подчинена никаким дальнейшим условиям: это вытекает из самого определения функции как «произвольного соответствия» (см. § 46).

Иначе обстоит дело, если заранее наложить на функцию $f(x)$ те или иные ограничения — выделить некоторый более узкий класс функций и рассматривать лишь функции, ему принадлежащие. Так, например, если говорить только об элементарных функциях (см. § 1), то можно было бы доказать, что, зная значения такой функции во всех точках некоторого промежутка сколь угодно малой длины, можно вычислить её значения в любой точке за пределами промежутка, лишь бы в этой точке функция не теряла смысла. Но вот другой пример: читатель согласится, как с фактом очевидным, что для определения линейной функции достаточно задать её значения всего лишь в двух точках; вообще для определения многочлена степени n достаточно указать его значения в $n+1$ различных точках.

В дальнейшем нас интересует, в какой степени значения функции в некотором множестве точек определяют её значения в точках, не принадлежащих этому множеству, если заранее известно, что функция $f(x)$ непрерывна.

---

[1]) См. стр. 192.

Легко понять, что из самого определения непрерывности вытекает следующее утверждение: если функция $f(x)$, заданная в некоторой окрестности точки $x=c$, непрерывна в самой этой точке, то значение её в этой точке может быть вычислено, раз известны её значения в точках некоторой последовательности $\{x_n\}$, имеющих пределом точку c.

В самом деле, в этом случае (см. § 42) должно быть:

$$f(c)=\lim f(x_n).$$

Вместо того чтобы задавать значения функции в точках последовательности $\{x_n\}$, имеющей пределом точку c, конечно, можно было бы также задать её значения во всех точках некоторого множества E — при единственном условии, чтобы из этого множества можно было выделить последовательность $\{x_n\}$, имеющую пределом точку c.

Особенно интересны с рассматриваемой точки зрения всюду плотные множества. Точечное множество E называется *всюду плотным*, если, каков бы ни был данный промежуток (α, β), $\alpha<\beta$, в этом промежутке содержится хоть одна точка E. Говорят также о множествах E, *всюду плотных в данном промежутке* I, если названному требованию удовлетворяет всякий промежуток (α, β), принадлежащий промежутку I.

Нам уже знакомы примеры всюду плотных множеств: таковы множества 1) рациональных чисел, 2) конечных десятичных дробей, 3) конечных бинарных дробей (чисел вида $\frac{m}{2^n}$, где m и n — целые).

Если множество E всюду плотно (в данном промежутке I), то, какова бы ни была точка c (из этого промежутка), всегда можно выделить такую последовательность точек $\{x_n\}$ из E, что точка c является её пределом:

$$x_n \to c. \tag{55}$$

Действительно, пусть $\{\varepsilon_n\}$ — убывающая последовательность положительных чисел и $\varepsilon_n \to 0$; из каждого промежутка $(c-\varepsilon_n, c+\varepsilon_n)$ $(n=1, 2, 3, \ldots)$ выберем точку x_n, принадлежащую E, и тогда будем иметь соотношение (55).

Заметим, между прочим, что последовательность $\{x_n\}$ может быть выбрана, если угодно, возрастающей или убывающей. Так, чтобы последовательность $\{x_n\}$ была возрастающей, достаточно взять x_1 из промежутка $(c-\varepsilon_1, c)$ и затем каждую следующую точку x_{n+1} выбирать внутри промежутка (x_n, c) $(n=2, 3, \ldots)$.

Сравнивая со сказанным выше, мы приходим к заключению: *если функция $f(x)$ непрерывна в некотором промежутке I, то по заданным её значениям в точках некоторого множества E, всюду плотного в промежутке I, можно вычислить её значения во всех точках I.*

Этот результат имеет ближайшее отношение к вопросу об определении показательной функции $f(x)=a^x$ $(a>1)$. Согласно формуле (69) из § 20:

$$f\left(\frac{p}{q}\right)=a^{\frac{p}{q}}=\sqrt[q]{a^p}; \tag{56}$$

показательная функция a^x определена (задана) во всех точках всюду плотного множества рациональных точек. Задача заключается в том, чтобы посредством операции перехода к пределу определить функцию a^x также и для иррациональных значений x. Пусть c — какое-нибудь иррациональное значение x. Действуя по предыдущей схеме, достаточно выделить такую последовательность рациональных чисел $\{r_n\}$, что $r_n\to c$, и тогда значение a^c должно определиться по формуле

$$f(c)=a^c=\lim_{n\to\infty} a^{r_n}. \tag{57}$$

Однако таким рассуждением нельзя удовлетвориться, так как нам заранее неизвестно, существует ли такая непрерывная функция $f(x)$, которая в рациональных точках $r=\frac{p}{q}$ принимает значения, указываемые равенством (56). При таких условиях: 1) подлежит доказательству существование предела в правой части (57), 2) необходимо убедиться, что этот предел не зависит от выбора последовательности рациональных чисел $\{r_n\}$.

Нисколько не облегчает положения то обстоятельство, что функция a^x является непрерывной по отношению ко множеству рациональных чисел E.

Рассмотрим следующие примеры. Пусть c — иррациональное число; функция $f_1(x)$, определённая при $x\neq c$ условиями

$$f_1(x)=\begin{cases} 0 \text{ при } x<c,\\ 1 \text{ при } x>c,\end{cases}$$

или функция $f_2(x)=\sin\frac{1}{x-c}$, определённая также при $x\neq c$, являются обе непрерывными по отношению ко множеству рациональных чисел; и тем не менее не представляется возможным приписать функции $f_1(x)$ или $f_2(x)$ такое значение в точке $x=c$, чтобы в этой точке они стали непрерывными.

Для того чтобы определить показательную функцию $f(x)=a^x$ $(a>1)$ в иррациональных точках, проще всего воспользоваться свойством её монотонности на множестве рациональных точек, т. е. свойством быть возрастающей, и, кроме того, — теоремой сложения (см. § 20).

Выберем возрастающую последовательность $\{r'_n\}$ и убывающую последовательность $\{r''_n\}$ рациональных чисел таким образом, чтобы было $r'_n\to c$, $r''_n\to c$; тогда из неравенств

$$r'_1<r'_2<\ldots<r'_n<\ldots<c<\ldots<r''_n<\ldots<r''_2<r''_1 \tag{58}$$

следуют неравенства

$$f(r'_1)<f(r'_2)<\ldots<f(r'_n)<\ldots<f(r''_n)<\ldots<f(r''_2)<f(r''_1). \tag{59}$$

Последовательность $\{f(r'_n)\}$ — возрастающая и ограниченная сверху; последовательность $\{f(r''_n)\}$ — убывающая и ограниченная снизу; значит, каждая из них имеет предел:

$$\left.\begin{aligned} f(r'_n) &\to C', \\ f(r''_n) &\to C''. \end{aligned}\right\} \qquad (60)$$

Вместе с тем, переходя к пределу в неравенстве

$$f(r'_n) < f(r''_n),$$

получаем

$$C' \leqslant C''.$$

Для доказательства того, что $C' = C''$, воспользуемся теоремой сложения:

$$a^{r''_n} = a^{r''_n - r'_n} \cdot a^{r'_n}. \qquad (61)$$

Так как, очевидно, $r''_n - r'_n \to 0$, то $a^{r''_n - r'_n} \to 1$ [1]), и переход к пределу в равенстве (61) даёт:

$$C'' = C'. \qquad (62)$$

Общее значение C' и C'' станем обозначать теперь через C (так что $f(r'_n) \to C$ и $f(r''_n) \to C$), и покажем, что если $\{\rho_n\}$ — какая угодно последовательность рациональных чисел, обладающая свойством $\rho_n \to c$, то непременно $f(\rho_n)$ стремится к C. Нужно убедиться, что, как бы мало ни было ε, при достаточно больших значениях n

$$|f(\rho_n) - C| < \varepsilon. \qquad (63)$$

Возьмём N настолько большим, чтобы выполнялись неравенства

$$C - \varepsilon < f(r'_N) \quad \text{и} \quad f(r''_N) < C + \varepsilon; \qquad (64)$$

тогда, так как c принадлежит промежутку (r'_N, r''_N), то при достаточно больших n мы получим

$$r'_N < \rho_n < r''_N,$$

и следовательно, по свойству возрастания $f(r)$,

$$f(r'_N) < f(\rho_n) < f(r''_N). \qquad (65)$$

Из сопоставления неравенств (64) и (65) затем следует неравенство (63).

Теперь значение функции $f(x) = a^x$ определено во всякой иррациональной точке $x = c$ по формуле (57). После этого непрерывность функции в этой точке уже не составляет проблемы: она следует из того обстоятельства, что при единственном условии $x_n \to c$ предел последовательности $\{f(x_n)\}$ существует и равен $C = f(c)$. Для какой угодно последовательности $\{x_n\}$ это доказывается так же, как только что было доказано для рациональной последова-

[1]) См. стр. 193.

тельности $\{\rho_n\}$, если принять во внимание, что добавление иррациональных точек не нарушает свойства показательной функции быть возрастающей.

Таким образом, показательная функция $f(x)=a^x$, определённая теперь при всех действительных значениях x, обладает свойствами непрерывности и монотонности во всём промежутке $-\infty<x<\infty$.

Наконец, остаётся в силе и «теорема сложения»

$$a^{x'+x''}=a^{x'}\cdot a^{x''}; \tag{66}$$

первоначально установленное (§ 20) для случая рациональных значений x' и x'', это соотношение со ссылкой на непрерывность обобщается на случай иррациональных значений. Пусть последовательности рациональных чисел $\{x'_n\}$ и $\{x''_n\}$ обладают свойствами $x'_n\to x'$ и $x''_n\to x''$; тогда достаточно перейти к пределу при $n\to\infty$ в равенстве

$$a^{x'_n+x''_n}=a^{x'_n}\cdot a^{x''_n}.$$

Укажем иной, более общий ход мыслей, позволяющий «по непрерывности» определять функцию $f(x)$ для всех значений x из некоторого промежутка I, если известны её значения в точках всюду плотного множества E.

Докажем теорему:

Если заданы значения функции $y=f(x)$ во всех точках множества E, всюду плотного в промежутке $I(a\leqslant x\leqslant b)$, причём

1° *на множестве E совокупность этих значений удовлетворяет требованию монотонности: с увеличением значения x увеличивается и значение $f(x)$,*

2° *заданное множество значений $f(x)$ в точках множества E принадлежит некоторому промежутку $K(A\leqslant y\leqslant B)$ и также всюду плотно в нём,*

то значения функции $f(x)$ в точках I, не принадлежащих E, могут быть определены таким образом, чтобы функция $f(x)$ была непрерывной во всём промежутке I.

Доказательство строится совершенно таким же образом, как в случае распространения понятия степени на случай иррационального показателя, только «множество рациональных точек» заменяется «множеством E», а роль произвольной иррациональной точки $x=c$ играет произвольная точка, принадлежащая промежутку I, но не множеству E. Заслуживает особого внимания лишь то, как доказать равенство (62), не прибегая к специальным свойствам показательной функции и используя зато условие 2°.

Допустим, что $C'<C''$; тогда вследствие 2° существует такое значение $x=r$ из множества E, что

$$C'<f(r)<C''. \tag{67}$$

Так как c не принадлежит E, то равенство $r=c$ невозможно. Пусть $r<c$; тогда при достаточно большом n получим $r<r'_n<c$ и, следовательно, $f(r)<f(r'_n)$; переходя к пределу, будем иметь $f(r)\leqslant C'$, что противоречит левому неравенству (67). Так же приходится отвергнуть и допущение $r>c$.

Итак, заключаем, что

$$C'=C''.$$

Пример 1. Чтобы доказанную теорему можно было, в частности, применить к функции $a^x\,(a>1)$, нужно только проверить, что множество

значений, принимаемых этой функцией в рациональных точках, всюду плотно. Пусть дан промежуток (α, β), причём $0 < \alpha < \beta$. Так как

$$\lim_{n \to \infty} \left(\frac{\beta}{\alpha}\right)^n = \infty,$$

то можно подобрать целое число n таким образом, что

$$\left(\frac{\beta}{\alpha}\right)^n > a,$$

откуда следует

$$q \equiv a^{\frac{1}{n}} < \frac{\beta}{\alpha}.$$

Легко понять, что хоть один из членов прогрессии

$$\ldots, q^{-2}, q^{-1}, 1, q, q^2, \ldots, q^n, \ldots$$

попадает в промежуток (α, β). Предположим, что таким членом будет q^m, так что

$$\alpha < q^m < \beta,$$

или

$$\alpha < a^{\frac{m}{n}} < \beta,$$

что и требовалось доказать.

Пример 2. В примере 2 § 46 множество K заданных значений функции $f(x)$ всюду плотно в промежутке (0, 1). В самом деле, обозначая через K_n множество значений, принимаемых функцией $f(x)$ в рациональных точках со знаменателем $\leqslant 2^n$, мы видим, что самый большой из промежутков, образованных на отрезке (0, 1) точками K_n имеет длину $\left(\frac{3}{4}\right)^n$. Пусть (α, β) — заданный промежуток $(0 \leqslant \alpha < \beta \leqslant 1)$; возьмём n удовлетворяющим условию $\left(\frac{3}{4}\right)^n < \beta - \alpha$; тогда на промежутке (α, β) непременно найдётся хоть одна точка множества K_n и, значит, множества K.

Пример 3. В примере 3 § 46 множество K заданных значений функции $f(x)$ также всюду плотно в промежутке (0, 1).

Предлагаем читателю доказать это в качестве упражнения.

Укажем, наконец, ещё третий способ построения показательной функции — посредством равномерного приближения многочленами.

Возвращаясь к многочленам

$$P_n(x) = \left(1 + \frac{x}{n}\right)^n \qquad (n = 1, 2, 3, \ldots),$$

рассмотренным в § 45, убедимся, что во всяком промежутке вида $(0, R)$ $(R > 0)$ последовательность $\{ P_n(x) \}$ равномерно стремится к некоторому пределу.

В самом деле, как в § 44, мы из сравнения коэффициентов в разложениях $P_n(x)$ и $P_{n+1}(x)$ по степеням x заключаем, что

$$0 < P_1(x) \leqslant P_2(x) \leqslant \ldots \leqslant P_n(x) \leqslant \ldots, \tag{68}$$

и вместе с тем при любом n

$$P_n(x) = 1 + \frac{n}{1}\left(\frac{x}{n}\right) + \frac{n(n-1)}{1\cdot 2}\left(\frac{x}{n}\right)^2 + \ldots + \frac{n(n-1)\ldots 1}{1\cdot 2\ldots n}\left(\frac{x}{n}\right)^n <$$

$$< 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} \leqslant$$

$$\leqslant 1 + \frac{R}{1!} + \frac{R^2}{2!} + \ldots + \frac{R^n}{n!}.$$

Пусть N — целое положительное число, $N > R$; тогда, далее, при $n > N$

$$P_n(x) \leqslant 1 + R + \frac{R^2}{2!} + \ldots + \frac{R^N}{N!} + \frac{R^N}{N!}\left\{\left(\frac{R}{N}\right) + \left(\frac{R}{N}\right)^2 + \ldots + \left(\frac{R}{N}\right)^{n-N}\right\} \leqslant$$

$$\leqslant 1 + R + \frac{R^2}{2!} + \ldots + \frac{R^N}{N!} + \frac{R^N}{N!}\frac{R}{N-R} \equiv M. \tag{69}$$

Из соотношений (68) и (69) следует, что последовательность $\{P_n(x)\}$ стремится к некоторому конечному пределу $f(x)$

$$P_n(x) \to f(x). \tag{70}$$

Следующие соображения показывают, что это стремление — равномерное. Так как разность $P_{m+1}(x) - P_m(x)$, будучи разложена по степеням x, имеет все коэффициенты положительные, то она лишь увеличится при замене x на R:

$$P_{m+1}(x) - P_m(x) \leqslant P_{m+1}(R) - P_m(R).$$

Просуммируем эти неравенства по букве m от n до $n+p-1$:

$$\sum_{m=n}^{n+p-1} [P_{m+1}(x) - P_m(x)] \leqslant \sum_{m=n}^{n+p-1} [P_{m+1}(R) - P_m(R)],$$

или

$$P_{n+p}(x) - P_n(x) \leqslant P_{n+p}(R) - P_n(R);$$

затем предельный переход $p \to \infty$ нам даёт:

$$f(x) - P_n(x) \leqslant f(R) - P_n(R). \tag{71}$$

Остаётся заметить, что при заданном ε и произвольном x из промежутка $[0, R]$ правая (не зависящая от x) часть неравенства (71), а значит, и левая может быть сделана меньше ε при достаточно большом n; итак, в этом промежутке

$$P_n(x) \rightrightarrows f(x).$$

Так как многочлены $P_n(x)$ — непрерывные функции, и сходимость $\{P_n(x)\}$ к $f(x)$ — равномерная, то на основании последней теоремы в § 48 функция $f(x)$ непрерывна в промежутке $[0, R]$. Число R здесь, однако, сколь угодно велико, поэтому функция $f(x)$ непрерывна в промежутке $[0, \infty)$.

С другой стороны, рассуждение, проведённое в § 45 (пример 1) при допущении, что значения x — рациональны (по нашему ходу мыслей показателю

степени x мы не имеем права давать иррациональных значений), показывает, что при всевозможных рациональных значениях x независимо от их знака имеет место равенство

$$f(x)=e^x.$$

Таким образом, функция $f(x)$, определяемая в промежутке $(-\infty, +\infty)$ соотношениями

$$\begin{cases} f(x)=\lim\limits_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n & \text{при } x\geqq 0, \\ f(x)=\dfrac{1}{f(-x)} & \text{при } x<0, \end{cases}$$

при рациональных значениях x равна e^x и вместе с тем при всех значениях x непрерывна.

Тем самым закончено построенне показательной функции e^x.

Примечание. Нетрудно понять, что, желая дать определение показательной функцин a^x при произвольном положительном основании a, достаточно было бы в предыдущей формуле заменить x через $x\ln a$:

$$a^x=\lim_{n\to\infty}\left(1+\frac{x\ln a}{n}\right)^n. \tag{72}$$

§ 52. Теорема Больцано и проблема существования однозначной обратной функции

Предположим, что функция $y=f(x)$ задана и непрерывна в замкнутом промежутке $I[a, b]$. Тогда по теореме Вейерштрасса (§ 47, III) она достигает в каких-то точках этого промежутка своего наименьшего значения A и своего наибольшего значения B. Исключая случай $A=B$ (когда функция сводится к постоянной), мы можем утверждать также на основании теоремы Больцано (§ 47, I), что в промежутке I функция $f(x)$ принимает — и, возможно, неоднократно — любое значение μ, заключённое между A и B. Итак, каково бы ни было значение y $(A\leqslant y\leqslant B)$, уравнение (относительно x)

$$f(x)=y \tag{73}$$

имеет по крайней мере один корень в промежутке $a\leqslant x\leqslant b$. Сопоставляя с каждым значением y из промежутка (A, B) все те значения x, которые являются корнями уравнения (73), мы получаем — вообще говоря, многозначную — обратную функцию

$$x=g(y), \tag{74}$$

которая задана в промежутке $A\leqslant y\leqslant B$ и в каждой его точке имеет по меньшей мере одно значение.

Весьма существенно уметь выделять те случаи, когда обратная функция $x=g(y)$ оказывается однозначной. Достаточные условия для этого даёт теорема:

Если функция $y=f(x)$ — непрерывная и возрастающая (убывающая) в промежутке $[a, b]$, то обратная функция $x=g(y)$ однозначна, возрастает (убывает) и, более того, непрерывна в промежутке $[A, B]$ [1]).

При этом $A=f(a)$, $B=f(b)$, если функция $f(x)$ — возрастающая, и $A=f(b)$, $B=f(a)$, если она — убывающая.

Обращаясь к доказательству, остановимся только на первом из этих двух предположений.

Существование решения уравнения (73) при $y=A$ или $y=B$ очевидно; при $A<y<B$ оно вытекает, как мы видели, из теоремы Больцано. Что решение — только одно, ясно почти сразу: если бы их было два, например x' и x'' $(x'<x'')$, то из равенств

$$f(x')=y \quad \text{и} \quad f(x'')=y$$

следовало бы

$$f(x')=f(x''),$$

тогда как по свойству возрастания $f(x)$ из неравенства $x'<x''$ должно вытекать, что

$$f(x')<f(x'').$$

Обозначим единственный корень уравнения (73) через $g(y)$, так что

$$f(g(y))=y \qquad (A<y<B).$$

Легко понять, что $g(y)$ является возрастающей функцией аргумента y. В самом деле, пусть

$$A\leqslant y'<y''\leqslant B. \tag{75}$$

Полагая $g(y')=x'$, $g(y'')=x''$, мы имеем:

$$f(x')=y',\ f(x'')=y''.$$

Отсюда понятно, что $x'<x''$; действительно, из противоположного допущения $x'\geqslant x''$, по свойству возрастания функции $f(x)$ следовало бы $f(x')\geqslant f(x'')$, т. е. $y'\geqslant y''$, а это противоречит предположению (75).

Остаётся обнаружить непрерывность функции $g(y)$. Рассмотрим хотя бы точку $y=d$, причём $A<d<B$. Так как функция $g(y)$ по доказанному — возрастающая, то существуют левый и правый пределы (см. § 42)

$$L'=g(d-0) \quad \text{и} \quad L''=g(d+0).$$

Притом из $y<d$ следует $g(y)<g(d)$; значит, переходя к пределу при $y\to d$, получаем: $L'\leqslant g(d)$. Аналогично, $g(d)\leqslant L''$. Итак,

$$L'\leqslant g(d)\leqslant L''.$$

---

[1]) Обратное предложение см. в конце § 62.

Нам нужно доказать, что

$$L' = g(d) = L''.$$

Допуская, напротив, например, что $L' < g(d)$, мы приходим к противоречию. В самом деле, выберем число ξ, удовлетворяющее неравенству

$$L' < \xi < g(d). \tag{76}$$

Очевидно, $a < \xi < b$, так как $g(d) < g(B) = b$ и $L' > g(A) = a$; следовательно, значению аргумента $x = \xi$ должно соответствовать некоторое значение функции $y = f(x)$. По свойству возрастания $f(x)$ из неравенства (76) следует:

$$f(L') < f(\xi) < f(g(d)). \tag{77}$$

Но $f(g(d)) = d$; с другой стороны, по свойству непрерывности $f(x)$ из соотношения $\lim\limits_{\substack{y \to d \\ y<d}} g(y) = L'$ следует:

$$\lim_{\substack{y \to d \\ y<d}} f(g(y)) = f(L')$$

и так как

$$\lim_{\substack{y \to d \\ y<d}} f(g(y)) = \lim_{\substack{y \to d \\ y<d}} y = d,$$

то

$$f(L') = d.$$

Итак, придавая неравенству (77) вид

$$d < f(\xi) < d,$$

убеждаемся в том, что оно противоречиво.

Заметим, что если рассматриваемый промежуток не замкнут с одного из концов (например, $a \leqslant x < b$, не исключая и случая $b = \infty$), то предыдущую теорему можно применить к замкнутому промежутку $a \leqslant x \leqslant b_1$ (где $b_1 < b$), с последующим предельным переходом $b_1 \to b$. Тогда оказывается, что заключение теоремы справедливо для незамкнутого промежутка $A \leqslant y < B$, причём $B = \lim\limits_{x \to b} f(x)$.

Пример 1. Функция $y = a^x$ $(a > 1)$ — непрерывная и возрастающая в промежутке $-\infty < x < +\infty$. Отсюда следует, что обратная функция $x = \log_a y$ — однозначная, возрастающая и непрерывная в промежутке $0 < x < \infty$.

Пример 2. Функция $y = x^n$ (n — целое положительное) — непрерывная и возрастающая в промежутке $0 \leqslant x < \infty$. Отсюда следует, что обратная функция $x = \sqrt[n]{y}$ (радикал — в арифметическом смысле) — возрастающая и непрерывная в промежутке $0 \leqslant y < \infty$.

Пример 3. Функция $y=\operatorname{tg} x$ — непрерывная и возрастающая в промежутке $-\frac{\pi}{2}<x<+\frac{\pi}{2}$. Значит, функция $y=\operatorname{arctg} x$ («главное значение» арктангенса) — возрастающая и непрерывная в промежутке $-\infty<y<+\infty$.

Пример 4. Функция $y=x^5+x$ — непрерывная и возрастающая в промежутке $0\leqslant x<\infty$. Отсюда следует, что x есть однозначная, возрастающая и непрерывная функция аргумента y в промежутке $0\leqslant y<\infty$. Эта функция, кстати сказать, не является элементарной в смысле § 1 (см. § 18, примечание).

§ 53. Функциональные уравнения и элементарные функции

Пусть $f(t)$ — функция, заданная в некотором промежутке I, и пусть x и y — какие-то произвольные значения аргумента t в этом промежутке; допустим также, что значение t, равное сумме $x+y$, также принадлежит тому же промежутку.

Положим ради краткости

$$X=f(x),\quad Y=f(y),\quad Z=f(x+y). \tag{78}$$

Допустим ещё, что из написанных трёх уравнений можно исключить две переменные x и y; другими словами, можно написать новое уравнение, скажем,

$$F(X,\ Y,\ Z)=0, \tag{79}$$

являющееся следствием уравнений (78) и вместе с тем не содержащее ни x, ни y. Подставляя сюда вместо X, Y, Z их значения, мы получим соотношение

$$F(f(x),\ f(y),\ f(x+y))=0, \tag{80}$$

справедливое при всех значениях x и y, подчинённых названным выше ограничениям.

Соотношение (80), принадлежащее к категории функциональных уравнений, носит название *теоремы сложения* функции $f(t)$ [1]).

Если бы вместо суммы $x+y$ мы взяли произведение xy, то, действуя как раньше, пришли бы к другому функциональному уравнению

$$\Phi(f(x),\ f(y),\ f(xy))=0, \tag{81}$$

которое носит название *теоремы умножения* функции $f(t)$.

Аналогичным образом, вводя вместо суммы $x+y$ или произведения xy некоторую произвольную функцию $\omega(x, y)$ переменных x и y, мы могли бы придти к функциональному уравнению вида

$$\Psi(f(x), f(y), f(\omega(x, y)))=0. \tag{82}$$

---

[1]) Необходимо объяснить, что соотношение (80) есть тождество относительно x и y; «уравнением» же его называют постольку, поскольку неизвестной считается функция f.

Но нам нет надобности итти по пути такого рода обобщений, и мы ограничимся рассмотрением теоремы сложения и теоремы умножения некоторых функций.

Рассмотрим примеры (в которых роль промежутка I будет играть вся числовая ось).

Пример 1. Функция $f(t)$ — линейная однородная:

$$f(t)=mt.$$

В таком случае уравнения (78) принимают вид

$$X=mx,\quad Y=my,\quad Z=m(x+y);$$

из них немедленно следует:

$$Z=X+Y$$

или

$$f(x+y)=f(x)+f(y). \tag{I}$$

Это и есть теорема сложения функции $f(t)=mt$.

Подобным же образом из уравнений

$$X=mx,\quad Y=my,\quad Z=mxy$$

получаем теорему умножения для тех же функций

$$mf(xy)=f(x)f(y). \tag{83}$$

Как видно, каждая из функций mt имеет свою теорему умножения, тогда как теорема сложения у них всех одна и та же.

Пример 2.

$$f(t)=a^t \qquad (a>0).$$

Из уравнений

$$X=a^x,\quad Y=a^y,\quad Z=a^{x+y}$$

следует уравнение

$$Z=XY,$$

т. е.

$$f(x+y)=f(x)f(y). \tag{II}$$

Это — теорема сложения для всех показательных функций вида a^t. Теорема умножения для функции a^t имела бы вид

$$\log_a f(xy)=\log_a f(x)\log_a f(y), \tag{84}$$

и следовательно, зависела бы от основания a.

Пример 3.

$$f(t)=\log_a t.$$

Теорема сложения оказывается зависящей от параметра a:

$$a^{f(x+y)}=a^{f(x)}+a^{f(y)},$$

тогда как теорема умножения от него не зависит:

$$f(xy)=f(x)+f(y). \tag{III}$$

Пример 4.

$$f(t)=t^{\alpha} \qquad (\alpha \neq 0).$$

И здесь теорема сложения зависит от параметра α:

$$[f(x+y)]^{\frac{1}{\alpha}}=[f(x)]^{\frac{1}{\alpha}}+[f(y)]^{\frac{1}{\alpha}},$$

тогда как теорема умножения от него не зависит:

$$f(xy)=f(x)f(y). \tag{IV}$$

Четыре уравнения (I), (II), (III) и (IV) в особенности привлекают наше внимание.

Посмотрим, исчерпывается ли уже известными нам формами совокупность их решений и в какой степени.

С некоторыми существенными оговорками ответ на этот вопрос оказывается утвердительным. Именно, справедливы следующие теоремы:

Теорема 1. *Если функция $f(t)$, заданная и непрерывная для всех значений t $(-\infty<t<+\infty)$, удовлетворяет уравнению* (I) *тождественно относительно x и y, то она имеет вид $f(t)=mt$, где m — постоянное число.*

Теорема 2. *Если функция $f(t)$, заданная и непрерывная для всех значений t $(-\infty<t<+\infty)$, удовлетворяет уравнению* (II) *тождественно относительно x и y, то она имеет вид $f(t)=a^t$, где a — неотрицательная постоянная.*

Теорема 3. *Если функция $f(t)$, заданная и непрерывная для всех положительных значений t, притом не сводящаяся к нулю, удовлетворяет уравнению* (III) *тождественно относительно всех положительных значений x и y, то она имеет вид $f(t)=\log_a t$, где a — положительная постоянная.*

Теорема 4. *Если функция $f(t)$, заданная и непрерывная для всех положительных значений t, притом не сводящаяся к нулю, удовлетворяет уравнению* (IV) *тождественно относительно всех положительных значений x и y, то она имеет вид $f(t)=t^{\alpha}$, где α — постоянное число.*

Доказательство теоремы 1. Посредством индукции по n из (I) легко выводится тождество

$$f(x_1+x_2+\ldots+x_n)=f(x_1)+f(x_2)+\ldots+f(x_n). \tag{85}$$

Полагая

$$x_1=x_2=\ldots=x_n=x,$$

получим

$$f(nx)=nf(x). \tag{86}$$

Далее, вводя обозначение

$$m=f(1),$$

при $x=1$ будем иметь (для всех целых положительных значений n):

$$f(n)=mn.$$

С другой стороны, заменяя в (86) n через q и полагая x равным $\frac{p}{q}$ (где p и q — целые положительные), приходим к соотношению

$$f(p)=qf\left(\frac{p}{q}\right),$$

откуда, если принять во внимание, что $f(p)=mp$, следует:

$$f\left(\frac{p}{q}\right)=m\frac{p}{q}. \tag{87}$$

Полагая затем в данном уравнении (I) $x=1$, $y=0$, получаем:

$$f(1)=f(1)+f(0),$$

так что

$$f(0)=0; \tag{88}$$

и наконец, подставляя в (I) $-x$ вместо y, будем иметь (при любом x)

$$0=f(x)+f(-x),$$

откуда

$$f(-x)=-f(x). \tag{89}$$

Сопоставляя (87), (88) и (89), мы можем сказать, что равенство

$$f(x)=mx \tag{90}$$

установлено для всюду плотного множества всех рациональных значений x. Отсюда, вследствие допущенной непрерывности $f(t)$, на основании теоремы, доказанной в § 51, и со ссылкой на непрерывность функции $f(t)=mt$ мы имеем возможность утверждать, что равенство (90) справедливо для всех действительных значений x.

Доказательство теоремы 2. Заменяя в уравнении (II) x и y через $\frac{t}{2}$, мы получаем:

$$f(t)=f^2\left(\frac{t}{2}\right)\geqslant 0.$$

Если бы функция $f(t)$ обращалась в нуль при некотором значении $t=x$, то вследствие (II) она была бы равна нулю также при $t=x+y$, где y совершенно произвольно, т. е. равнялась бы нулю тождественно.

Оставляя в стороне это предположение, соответствующее случаю $a=0$, мы должны считать, что при всех значениях t

$$f(t)>0.$$

Пусть $f(1)=a>0$. Полагая в таком случае

$$f(t)=a^{\varphi(t)}, \quad \varphi(t)=\log_a f(t),$$

мы видим, что уравнению (II) можно придать вид

$$a^{\varphi(x+y)}=a^{\varphi(x)}\,a^{\varphi(y)},$$

т. е.

$$\varphi(x+y)=\varphi(x)+\varphi(y).$$

Это — уравнение типа (I); и из доказательства теоремы 1 следует, что

$$\varphi(t)=mt$$

или

$$f(t)=a^{mt}.$$

Подставив значение $t=1$, убеждаемся, что $m=1$, так что $f(t)=a^t$.

Доказательство теоремы 4. Поскольку x и y предполагаются произвольными положительными числами, можно положить

$$x=10^{\xi}, \quad \xi=\lg x,$$
$$y=10^{\eta}, \quad \eta=\lg y,$$

и тогда уравнение (IV) примет вид

$$f(10^{\xi+\eta})=f(10^{\xi})f(10^{\eta}). \tag{IV'}$$

Вводя ещё новую неизвестную функцию

$$F(\tau)\equiv f(10^{\tau}), \tag{91}$$

мы придадим (IV) вид

$$F(\xi+\eta)=F(\xi)F(\eta).$$

Так как ξ и η — произвольные действительные числа, и функция $F(\tau)$ непрерывна на всей оси (по теореме о непрерывности сложной функции, вследствие непрерывности функций $f(t)$ и 10^{τ}), то на основании теоремы 2 получим:

$$F(\tau)\equiv a^{\tau} \qquad (a\geqslant 0).$$

В таком случае, заменяя в последнем тождестве τ через $\lg t$, мы получаем окончательно:

$$f(t)\equiv a^{\lg t}\equiv\begin{cases}0, & \text{если } a=0,\\ t^{\alpha} \text{ (где } \alpha=\lg a), & \text{если } a>0.\end{cases}$$

Доказательство теоремы 3. Уравнение (III) можно переписать в виде

$$10^{f(xy)}=10^{f(x)}\,10^{f(y)}.$$

Вводя функцию

$$G(t)\equiv 10^{f(t)}, \tag{92}$$

придадим ему вид

$$G(xy)=G(x)\,G(y). \tag{III'}$$

По теореме 4 из (III') следует, что или $G(t) \equiv 0$ или $G(t) \equiv t^{\alpha}$. Так как показательная функция в нуль не обращается, то приходится рассматривать лишь вторую возможность. Но если $G(t) = t^{\alpha}$, то при допущении $\alpha = 0$ мы получаем $G(t) \equiv 1$, $f(t) \equiv 0$; оставляя эту возможность в стороне, приходим к заключению, что

$$f(t) \equiv \alpha \lg t = \log_a t \qquad (\text{где } a = 10^{\frac{1}{\alpha}}).$$

Всё изложенное выше показывает, что функциональные уравнения (I—IV) являются характеристическими соответственно для функций mt, a^t, $\log_a t$ и t^{α}; другими словами, эти уравнения могут служить в качестве определения этих функций — при дополнительном требовании непрерывности.

Примечание 1. Если вместо непрерывности наложить на функции $f(t)$ требование дифференцируемости (см. стр. 309), то тем более заключевия теорем 1—4 остаются в силе. Доказательства теорем при таких условиях значительно упрощаются. Например, теорема 1 может быть доказана следующим образом. Дифференцируя тождество (I) по переменной x, мы получаем:

$$f'(x+y) = f'(x),$$

и так как y здесь — произвольное, то отсюда вытекает, что $f'(x)$ сводится к постоянной:

$$f'(x) = m.$$

Но тогда интегрирование даёт:

$$f(x) = mx + C,$$

где C — новая постоянная. Подставляя найденное выражение для $f(x)$ снова в уравнение (I), видим, что при всех значениях x и y

$$m(x+y) + C = (mx + C) + (my + C),$$

откуда следует, что $C = 0$ и что $f(x) = mx$.

Среди функций, которые мы смогли определить с помощью уравнений (I—IV), не фигурируют тригонометрические, а также им обратные.

Читатель, который пожелал бы составить теорему сложения, например, для функции $f(t) = \sin t$, пришёл бы к известной из тригонометрии формуле для «синуса суммы»; но в ней косинусы были бы выражены через синусы, так что формула имела бы следующий, сравнительно сложный вид:

$$f(x+y) = f(x) \cdot \sqrt{1 - f^2(y)} + f(y) \cdot \sqrt{1 - f^2(x)}$$

(можно было бы «избавиться от иррациональности»).

Уместно, впрочем, высказать следующие соображения. В теории функций комплексного переменного устанавливается, что тригонометрические функции выражаются, с привлечением мнимой единицы, через показательные, например,

$$\cos t \equiv \frac{e^{it} + e^{-it}}{2}, \quad \sin t \equiv \frac{e^{it} - e^{-it}}{2i}, \tag{93}$$

теряя при этом, так сказать, право на самостоятельное существование; точно так же обратные тригонометрические функции выражаются через логарифмы и радикалы (см. стр. 511). Таким образом, при перечислении

основных операций, служащих для получения элементарных функций, к четырём арифметическим действиям (сложение, вычитание, умножение, деление) достаточно прибавить всего лишь потенцирование, логарифмирование и возведение в произвольную степень [1]).

Принимая это во внимание, мы приходим к мысли о справедливости следующего утверждения:

Все элементарные функции составляются из независимой переменной в результате повторного применения конечного числа арифметических действий и тех операций, которые, будучи обозначены символом f, являются решениями функциональных уравнений (I—IV).

Отсюда выясняется то значение (конечно, теоретическое), которое имеют уравнения (I—IV) для элементарной математики.

Примечание 2. Из теорем 1—4 предыдущее утверждение, строго говоря, не вытекает по той причине, что при составлении тригонометрических функций из показательной согласно формулам (93) приходится прибегать к операции $f(t) = it$, тогда как в решении $f(t) = mt$ уравнения (I) постоянная m предполагается действительной. Чтобы обосновать наше утверждение, надо перенести теоремы 1—4 в комплексную область. Естественно, что при таком перенесении требование непрерывности заменится требованием регулярности искомой функции $f(t)$. Так, формулируя теорему 1′, аналогичную теореме 1, нужно предполагать функцию f регулярной во всей плоскости и допускать, что соотношение (I) выполняется при всех комплексных значениях x и y. Так как из регулярности следует дифференцируемость, то доказательство было бы построено далее так, как указано в примечании 1, но при этом постоянная m могла бы оказаться произвольной комплексной.

Рекомендуем читателю, по прочтении статьи, «Элементарные функции комплексного переменного» (стр. 493—552), попытаться сформулировать, как эту, так и остальные три теоремы, касающиеся решения уравнений (I—IV) в комплексной плоскости, а также восстановить детали доказательств.

---

[1]) Напомним, кстати, что возведение в произвольную степень конструируется из логарифмирования, умножения на постоянное число и потенцирования (см. стр. 93).

ГЛАВА V

ОБЩЕЕ ПОНЯТИЕ ФУНКЦИИ

§ 54. Соответствие между множествами

В § 1 было упомянуто, а в § 46 подробно рассмотрено определение функции как соответствия между двумя числовыми множествами, установленного на основе совершенно произвольного правила; как ни общо это определение, всё же в другом отношении оно является очень узким и допускает значительное расширение. В силу этого определения функция

$$y = f(x)$$

сопоставляет с каждым числом (с числовым значением переменной x из некоторого промежутка) некоторое число (соответствующее значение y).

В более общем случае можно говорить о множествах, составленных из элементов какой угодно природы. Именно, общее определение функции, имеющее громадное значение в современной математике, формулируется следующим образом.

Дано некоторое множество объектов, которое обозначим через $\mathcal{X}$; сами объекты, как элементы этого множества, условимся обозначать буквой X («переменное» во множестве $\mathcal{X}$), снабжая её различными значками, если речь идёт об отдельных элементах множества $\mathcal{X}$. Пусть также задано некоторое множество объектов, которое назовём $\mathcal{Y}$: оно может содержать полностью или частично элементы множества $\mathcal{X}$, но может также состоять сплошь из элементов, не принадлежащих множеству $\mathcal{X}$; элементы этого множества $\mathcal{Y}$ мы будем обозначать буквой Y («переменное» во множестве $\mathcal{Y}$), также снабжая её различными значками, если хотим выделить отдельные элементы.

Говорят, что при этих условиях *«переменное Y есть функция переменного X»* и пишут

$$Y = f(X)\ ^{1}), \qquad (1)$$

[1]) Лишь ради простоты письма приходится пользоваться знаком обыкновенного равенства там, где обязателен был бы знак тождества ($\equiv$): ведь элементы Y и $f(X)$ могут не быть числами.

если указано некоторое правило, сопоставляющее с каждым элементом X из множества $\mathcal{X}$ некоторый элемент Y из множества $\mathcal{Y}$. При этом, вообще говоря, не требуется, чтобы каждый элемент множества $\mathcal{Y}$ непременно был поставлен в соответствие некоторому элементу множества $\mathcal{X}$, и не требуется также, чтобы с различными элементами из $\mathcal{X}$ были сопоставлены различные же элементы из $\mathcal{Y}$. Напротив, может случиться, что во множестве $\mathcal{Y}$ останутся «свободные», или «лишние» элементы, не сопоставленные ни с какими элементами из $\mathcal{X}$, или что с различными элементами X' и X'' из $\mathcal{X}$ будет сопоставлен один и тот же элемент из $\mathcal{Y}$.

При указанных обстоятельствах говорят также об *отображении* каждого элемента X множества $\mathcal{X}$ на сопоставляемый с ним элемент Y множества $\mathcal{Y}$ и об отображении совокупности всех элементов множества $\mathcal{X}$ на совокупность всех сопоставляемых с ними элементов множества $\mathcal{Y}$.

Описанное отображение общего типа называется *однозначным* — в противоположность *многозначному* отображению, при котором с каждым элементом из $\mathcal{X}$ сопоставляется один или несколько элементов из $\mathcal{Y}$.

В дальнейшем речь будет итти лишь об однозначных отображениях.

Отображение множества $\mathcal{X}$ на множество $\mathcal{Y}$ называется *взаимно однозначным*, если выполнено требование, чтобы каждый элемент из $\mathcal{Y}$ был сопоставлен с некоторым элементом из $\mathcal{X}$ и чтобы с различными элементами из $\mathcal{X}$ сопоставлялись различные же элементы из $\mathcal{Y}$; другими словами, если из соотношения

$$f(X') = f(X'')$$

следует соотношение

$$X' = X''.$$

В этом случае, обратно, отображение (1) с каждым элементом Y из $\mathcal{Y}$ сопоставляет некоторый элемент X из $\mathcal{X}$; именно, единственный тот, с которым сопоставлен элемент Y; другими словами, существует обратное отображение Y на $\mathcal{X}$:

$$X = g(Y). \tag{2}$$

В частности, если $\mathcal{X}$ и $\mathcal{Y}$ — некоторые множества действительных чисел, мы получаем определение функции одного переменного на множестве E, рассмотренное в § 46. Чаще всего E — промежуток, замкнутый или открытый.

В следующих параграфах мы рассмотрим ряд примеров отображений, причём роль элементов множеств $\mathcal{X}$ и $\mathcal{Y}$ будут играть не только числа, но и пары чисел, тройки чисел и т. д., вообще системы чисел, а также функции определённого класса, заранее заданные в некоторой фиксированной области.

При этом будут указываться и некоторые геометрические или механические (кинематические) интерпретации, служащие, с одной стороны (как и «графики» функций в простейшем случае), целям наглядности, но вместе с тем указывающие и конкретные «реализации» общей схемы.

Однако обзор будет довольно беглым — не только по ограниченности места, предоставленного настоящей статье, но и по той причине, что нельзя рассчитывать, чтобы возникающие концепции могли найти постоянное употребление в средней школе.

В заключение настоящего параграфа приведём несколько примеров «функциональных соответствий» или «отображений», где каждое из множеств $\mathcal{X}$ и $\mathcal{Y}$ содержит конечное число элементов.

В таком случае число возможных отображений — тоже конечное: оно равно n^m, если $\mathcal{X}$ содержит m, а $\mathcal{Y}$ — n элементов [1]).

1. $\mathcal{X}$ — множество пальцев левой руки, $\mathcal{Y}$ — множество пальцев правой руки.

Всего возможно $5^5 = 3125$ соответствий, из них взаимно однозначных $P_5 = 5! = 120$.

2. $\mathcal{X}$ — множество учеников в классе (30), $\mathcal{Y}$ — множество различных баллов (1—5). В результате письменной работы устанавливается одно из соответствий. Всего возможно 5^{30} соответствий.

3. $\mathcal{X}$ — множество баллов (1—5), $\mathcal{Y}$ — множество учеников в классе (30). Из района требуют прислать образчики работ — по одной на каждый из баллов. Всего возможно 30^5 соответствий.

4. Каждый из трёх товарищей может по желанию пойти или не пойти на экскурсию. $\mathcal{X}$ — множество товарищей, $\mathcal{Y}$ — множество, состоящее из двух значков: $+$ («пойдёт») и $-$ («не пойдёт»). Возможно $2^3 = 8$ соответствий.

5. m различных предметов нужно разложить в n ящиках. $\mathcal{X}$ — множество предметов, $\mathcal{Y}$ — множество ящиков. Разложить можно n^m способами.

§ 55. Геометрические образы в многомерных пространствах

Предположим, что $\mathcal{X}$ — множество всех чисел z (или чисел z, содержащихся в некотором промежутке I), $\mathcal{Y}$ — множество всевозможных пар чисел (x, y). Каждому значению z соответствует некоторая пара (x, y) и, значит, некоторое значение x и некоторое значение y; итак, x и y являются функциями (в обычном смысле) независимой переменной z:

$$\left.\begin{aligned} x &= f(z), \\ y &= g(z). \end{aligned}\right\} \tag{3}$$

Уясним себе геометрический смысл такого соответствия, воспользовавшись координатной системой $Oxyz$ в пространстве. Соот-

[1]) С каждым из элементов множества $\mathcal{X}$ может быть сопоставлен любой из элементов множества $\mathcal{Y}$. Если $m = 1$, то, очевидно, имеется n возможных отображений; если $m = 2$, то получается n^2 возможных отображений, так как каждое из n возможных отображений первого элемента может «скомбинироваться» с каждым из n возможных отображений для второго элемента; и т. д. — по индукции.

ношения (3) говорят о том, что в каждой плоскости, параллельной плоскости Oxy, выделяется одна точка: совокупность этих точек образует [1]) некоторую кривую в пространстве.

Пусть читатель рассмотрит более внимательно, например, винтовую линию

$$\left.\begin{aligned} x &= \cos z, \\ y &= \sin z \end{aligned}\right\} \qquad (4)$$

(рис. 90) и разберётся в том, как при изменении z от 0 до 2π точка с координатами (x, y, z) совершает один «обход» вокруг оси Oz.

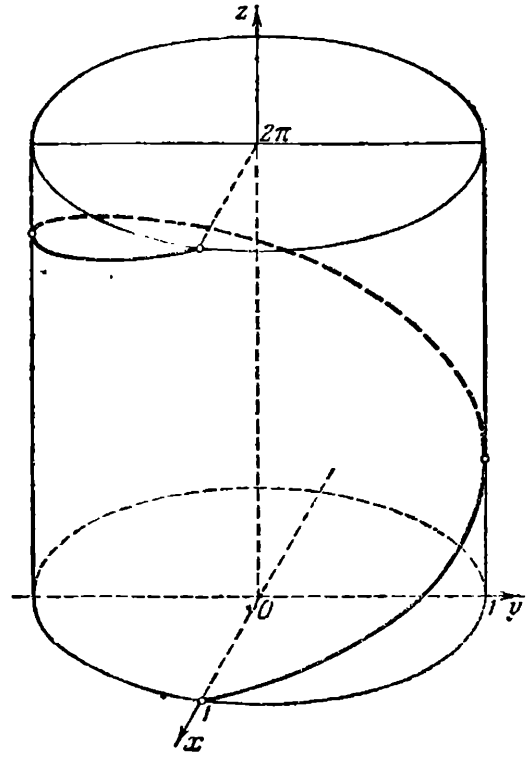


Рис. 90.

Предположим теперь, что $\mathscr{X}$ — множество (D) всевозможных пар чисел (x, y) или, другими словами, $\mathscr{X}$ — множество всех точек координатной плоскости Oxy; можно также предположить, что это множество (D) составлено из всех точек плоскости Oxy, лежащих внутри некоторой замкнутой кривой или по одну сторону некоторой разомкнутой кривой, идущей из бесконечности в бесконечность. Что касается $\mathscr{Y}$, то допустим, что это — множество всех чисел z.

Таким образом, с каждой парой чисел (x, y) из (D) сопоставлено некоторое число z, которое является функцией двух переменных x и y:

$$z = f(x, y). \qquad (5)$$

Написанное соотношение говорит о том, что в пространстве $Oxyz$ на всякой прямой, параллельной оси Oz и проходящей через какую-нибудь точку множества (D) на плоскости Oxy, выделяется единственная точка; совокупность этих точек (при соблюдении условий непрерывности) образует некоторую поверхность в пространстве.

Рассмотрим, например, уравнение

$$z = \sin \sqrt{x^2 + y^2}. \qquad (6)$$

---

[1]) При соблюдении условия непрерывности функций $f(z)$ и $g(z)$.

Принимая во внимание, что выражение $\sqrt{x^2+y^2}$ представляет расстояние точки (x, y) в плоскости Oxy от начала O, можно заключить в данном примере, что уравнение выделяет на прямой, проходящей через точку (x, y) и параллельной оси Oz, точку с координатой z, равной синусу от упомянутого расстояния; нетрудно отдать себе отчёт в том, что совокупность точек (x, y, z) в пространстве образует «гофрированную» поверхность, которая получается при вращении «полусинусоиды» $z=\sin x$ $(0\leqslant x<\infty)$, взятой в плоскости Oxz, относительно оси Oz (рис. 91).

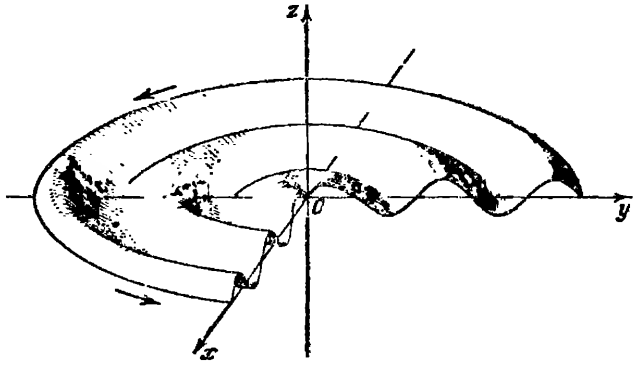


Рис. 91.

Другой пример поверхности, которую можно рекомендовать вниманию читателя, — «винтовая поверхность» с осью Oz

$$z=\operatorname{Arctg}\frac{y}{x}. \tag{7}$$

Перейдём от случаев, когда общее число рассматриваемых переменных равнялось трём, к случаю, когда оно больше трёх. Если число переменных равно четырём, то, обозначая их, скажем, через x, y, z и u, можно выделить три существенно различных случая.

1) Каждая из трёх переменных является функцией четвёртой, например

$$\left.\begin{aligned} x&=f(u),\\ y&=g(u),\\ z&=h(u); \end{aligned}\right\} \tag{8}$$

здесь $\mathscr{X}$ состоит из множества чисел u, $\mathscr{Y}$ — из множества троек чисел (x, y, z).

2) Каждая из двух переменных является функцией остальных двух, например

$$\left.\begin{aligned} z&=f(x, y),\\ u&=f(x, y); \end{aligned}\right\} \tag{9}$$

здесь $\mathscr{X}$ состоит из множества пар чисел (x, y), $\mathscr{Y}$ — также из множества пар (z, u).

3) Одна из переменных является функцией трёх остальных, например

$$u = f(x, y, z); \tag{10}$$

здесь $\mathscr{X}$ состоит из множества троек чисел (x, y, z), $\mathscr{Y}$ — из множества чисел u.

При этом во всех случаях множество $\mathscr{X}$ может быть подчинено тем или иным ограничениям.

Упомянем, наконец, гораздо более общий типический случай, когда множество $\mathscr{X}$ состоит из систем p чисел, «независимых» переменных $(x_1, x_2, \ldots, x_p)$, а множество $\mathscr{Y}$ — из систем q чисел, «зависимых» переменных $(y_1, y_2, \ldots, y_q)$:

$$\left.\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_p), \\ y_2 &= f_2(x_1, x_2, \ldots, x_p), \\ &\cdots\cdots\cdots\cdots \\ y_q &= f_q(x_1, x_2, \ldots, x_p). \end{aligned}\right\} \tag{11}$$

Что касается геометрического представления, то, если следовать общему принципу, в этом случае нужно было бы обратиться к n-мерному пространству $Ox_1x_2 \ldots x_py_1y_2 \ldots y_q$, где $n = p + q$, и тогда система функциональных соотношений (11) могла бы быть истолкована как определяющая некоторый p-мерный геометрический образ, понимая под таковым совокупность «точек»

$$(x_1, x_2, \ldots, x_p, y_1, y_2, \ldots, y_q),$$

удовлетворяющих одновременно всем соотношениям (11).

Уравнения, определяющие в своей совокупности некоторую протяжённость [1]), могут и не быть «решены» относительно зависимых переменных, и тогда в случае, если протяжённость задана «неявно», не обязательно каждой системе $(x_1, x_2, \ldots, x_p)$ соответствует именно одна система $(y_1, y_2, \ldots, y_q)$: может возникнуть «многозначность».

В качестве примера укажем уравнение «сферы» в n-мерном пространстве

$$x_1^2 + x_2^2 + \ldots + x_n^2 = R^2.$$

Решая его, например, относительно x_n $(\equiv y_1)$, получаем при условии $x_1^2 + x_2^2 + \ldots + x_{n-1}^2 \leqslant R^2$ «двузначную» функцию

$$y_1 = \pm\sqrt{R^2 - (x_1^2 + x_2^2 + \ldots + x_{n-1}^2)}.$$

Здесь $p = n - 1$, $q = 1$.

---

[1]) Термин «протяжённость» охватывает всевозможные типы геометрических образов в n-мерном пространстве ($n \geqslant 4$). Счёт «измерений» ведётся по числу независимых переменных. Название «кривой» обычно сохраняется за одномерными протяжённостями, название «поверхности» — за $(n-1)$-мерными.

Заметим ещё (ограничиваясь хотя бы случаем трёхмерного пространства), что одна и та же протяжённость может быть задана различными эквивалентными между собой комбинациями уравнений. Например, винтовая линия задана парой уравнений (4) как пересечение двух «синусоидальных цилиндрических поверхностей» с образующими, соответственно параллельными осям Oy и Ox; но её же можно задать парой эквивалентных уравнений

$$\left.\begin{aligned} x^2+y^2&=1,\\ z&=\operatorname{Arctg}\frac{y}{x},\end{aligned}\right\} \tag{12}$$

как пересечение кругового цилиндра с осью Oz и «винтозой поверхности» — с той же осью.

§ 56. Пространственные отображения

Пытаясь побудить работать не только мысль, но и воображение, мы постоянно прибегаем к «геометрическим представлениям», или «интерпретациям» аналитических объектов или взаимоотношений. Однако мы свободны в выборе этих интерпретаций: с одними и теми же аналитическими фактами могут быть связываемы различные геометрические представления.

Исходя из этой общей идеи, можно функциональной зависимости между двумя переменными величинами (числами или системами чисел) дать иную, пожалуй более простую и естественную, интерпретацию.

В случае функциональной зависимости

$$y=f(x) \tag{13}$$

мы прибегали к координатной плоскости Oxy и, толкуя x как абсциссу, y — как ординату точки этой плоскости, сопоставляли с уравнением (13) некоторый график — кривую линию, лежащую в упомянутой плоскости.

Но ничто не помешало бы вместо плоскости Oxy рассматривать две независимые между собой, как угодно расположенные координатные оси Ox и Oy [1]); при этом зависимость (13) толкуется в том смысле, что *с каждой точкой на оси Ox сопоставляется некоторая точка на оси Oy.*

На рис. 92 намечен пример функции $y=\frac{1}{10}x^2$, причём соответствия между точками обозначены стрелочками.

Точно так же можно считать, что зависимость вида (3) *сопоставляет с каждой точкой оси Oz* (или некоторого промежутка

---

[1]) Совпадение букв O не означает совмещения начал на двух осях.

на этой оси) *некоторую точку независимой* (произвольно расположенной) *плоскости Oxy.*

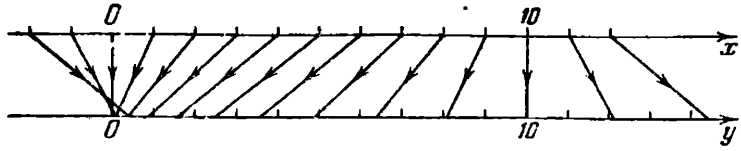


Рис. 92.

Рисунок 93 соответствует примеру (4); в данном случае промежуток $0 \leqslant z < 2\pi$, и вообще $2k\pi \leqslant z < 2(k+1)\pi$, отображается на окружность $x^2 + y^2 = 1$.

Зависимости же вида (5), напротив, интерпретируются как *соответствие между точками плоскости Oxy* (или некоторой области (D) в этой плоскости) *и точками независимой оси Oz.* Рисунок 94 иллюстрирует пример, заданный уравнением (6). Здесь с любой точкой одной и той же окружности в плоскости *Oxy* сопоставляется некоторая одна и та же точка оси *Oz*, и именно, например, часть плоскости *Oxy*, находящаяся в окружности $x^2 + y^2 = \frac{\pi^2}{4}$, отображается таким образом на промежуток $0 \leqslant z < 1$.

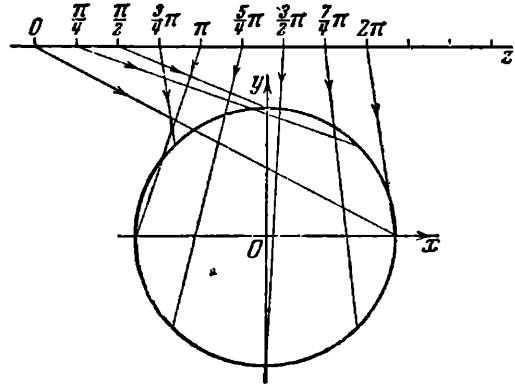


Рис. 93.

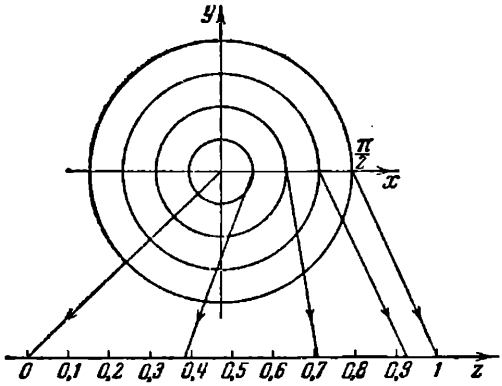


Рис. 94.

Подобным же образом в случае четырёх переменных зависимости типа (8), (9) и (10) интерпретируются соответственно как отображение прямой на пространство, плоскости на плоскость и пространства на прямую.

Вообще *зависимость типа* (11) *интерпретируется как отображение «p-мерного пространства»* $Ox_1x_2 \ldots x_p$ *на «q-мерное пространство»* $Oy_1y_2 \ldots y_q$.

Приведём несколько примеров такого рода отображений, не обязательно при этом воспроизводя формулы — потому ли, что

воспроизвести их не стоит труда или потому, напротив, что воспроизведение их, точное или приближённое, связано с более или менее значительными принципиальными или практическими затруднениями.

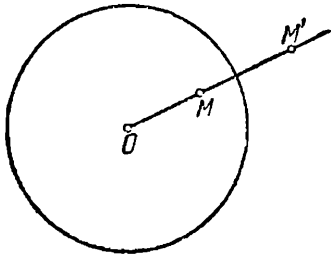


Рис. 95.

1. Симметрические отображения, параллельные перенесения, вращения, растяжения или сжатия (см. многочисленные примеры в главе I) представляют собой частные случаи отображения плоскости на плоскость; они взаимно однозначны; все они переносятся и на трёхмерное пространство.

2. Рассмотрим следующее специальное отображение плоскости на плоскость (в данном случае отображение плоскости самой на себя). Дана окружность с центром O и радиусом R (рис. 95). С точкой M сопоставляется точка M', лежащая на луче OM и обладающая свойством

$$OM \cdot OM' = R^2.$$

Это отображение, как легко понять, — также взаимно однозначное и также переносится на случай пространства. Оно носит название *инверсии* и обладает замечательными свойствами.

Если положим $R = 1$ и поместим центр окружности в начало координат, то уравнение окружности примет вид

$$x^2 + y^2 = 1.$$

Обозначая координаты точки M через x, y, а координаты точки M' — через x', y', получим уравнения инверсии в аналитической форме:

$$\begin{cases} x' = \dfrac{x}{x^2+y^2}, \\ y' = \dfrac{y}{x^2+y^2}, \end{cases} \quad \text{или} \quad \begin{cases} x = \dfrac{x'}{x'^2+y'^2}, \\ y = \dfrac{y'}{x'^2+y'^2}. \end{cases}$$

3. Вообразим идеальный сосуд (V), наполненный жидкостью в состоянии покоя; взболтаем жидкость и дождёмся момента, когда она снова вернётся в состояние покоя. Тогда, сопоставляя с положением каждой частицы жидкости до взбалтывания положение той же частицы после взбалтывания, получим отображение части пространства (V) на ту же самую часть пространства.

4. Пусть независимое переменное t истолковывается как время с ограничением $t_1 \leqslant t \leqslant t_2$; роль же зависимого переменного играет положение точки, движущейся а) по прямой Ox, б) на плоскости Oxy, в) в пространстве $Oxyz$. Тогда движение точки определяется уравнениями или системами уравнений вида

$$\text{а) } x = f(t); \quad \text{б) } \begin{cases} x = f(t), \\ y = g(t); \end{cases} \quad \text{в) } \begin{cases} x = f(t), \\ y = g(t), \\ z = h(t). \end{cases} \tag{14}$$

Тем самым «отрезок времени» $t_1 \leqslant t \leqslant t_2$ отображается на прямую, на плоскость или на пространство. Взаимной однозначности здесь может не быть, так как в различные моменты времени точка может занимать одно и то же положение.

Частные примеры:

1) $\begin{cases} x = \cos t, \\ y = \sin t \end{cases}$ (равномерное вращение точки по кругу $x^2 + y^2 = 1$);

2) $\begin{cases} x = \operatorname{ch} t, \\ y = \operatorname{sh} t \end{cases}$ (см. § 21: движение точки по гиперболе $x^2 - y^2 = 1$) [1]);

3) $\begin{cases} x = \cos t, \\ y = \sin t, \\ z = t \end{cases}$ (движение точки по винтовой линии в пространстве).

5. Пусть u — какая угодно величина, значения которой изображаются точками оси Ou. Величина u пусть будет произвольной «функцией точки»; например, u может быть температурой, давлением и т. п. Тогда уравнения вида

$$\text{а) } u = f(x), \quad \text{б) } u = f(x, y), \quad \text{в) } u = f(x, y, z),$$

указывающие распределение величины u а) на прямой Ox, б) на плоскости Oxy, в) в пространстве $Oxyz$ (или в какой-нибудь их части), вместе с тем определяют отображение прямой Ox, плоскости Oxy или пространства $Oxyz$ на прямую Ou. Совокупность точек, в которых u имеет заданное значение (т. е. отображаемых на одну и ту же точку оси Ou), носит название *кривой* (в случае б)) или *поверхности* (в случае в)) *уровня*. Если u — температура, это — *изотерма*; если u — давление, то *изобара* и т. п.

Следующие примеры — несколько иного характера: они связаны с отображением кривой на кривую или поверхности на поверхность.

6. Одна окружность (c) находится внутри другой (C) и дана ещё точка O внутри c. С точкой m на окружности (c) сопоставляется та точка M окружности (C), которая лежит на луче Om.

То же можно сделать с поверхностями сфер в пространстве.

7. Обвернём без перекрытий цилиндр листом бумаги и поставим во взаимное соответствие ту точку поверхности цилиндра и ту точку листа бумаги, которые при обворачивании совмещаются. Затем, развернув лист бумаги на плоскость, получим отображение цилиндра на плоскость, при котором длины взаимно соответствующих кривых одни и те же. То же можно сделать с конусом.

Говорят в таких случаях о «развёртывании» цилиндра или конуса на плоскость.

---

[1]) Отсюда наименование «гиперболические функции».

8. Но того же нельзя сделать с шаром. Тем не менее необходимо (например, в картографии) каким-то образом отображать поверхность земного шара на плоскость; это делается обычно по особым правилам, выражающимся аналитическими формулами.

Самое простое правило заключается в том, что долгота v и широта u точки земной поверхности интерпретируются как прямоугольные координаты — абсцисса и ордината. Читатель может, руководствуясь этим принципом, попытаться наметить на клетчатой бумаге карту Европы.

§ 57. Метрические пространства

В § 38 было дано определение предела последовательности чисел, связанное теснейшим образом с понятием окрестности точки (на числовой оси); при этом под *окрестностью* понимается обыкновенно совокупность точек, расстояние (отклонение) которых от рассматриваемой точки меньше данного числа.

Аналогичным образом можно определить предел последовательности элементов любого множества, лишь бы предварительно было установлено понятие «расстояния» между всякими двумя элементами этого множества. Другими словами, должно быть «введено мероопределение», или «установлена метрика».

Если речь идёт о числах, например x' и x'', то под расстоянием между двумя числами естественно понимать (см. § 36) абсолютную величину их разности $|x' - x''|$, что как раз соответствует расстоянию в обычном геометрическом смысле между соответствующими точками M' и M'' числовой прямой:

$$M'M'' = |x' - x''|. \tag{15}$$

Если речь идёт о парах чисел, скажем (x', y') и (x'', y''), то, прибегая к обычному их представлению в виде точек координатной (декартовой) плоскости Oxy, можно под «расстоянием» понимать обычное геометрическое расстояние между соответствующими точками $M'(x', y')$ и $M''(x'', y'')$: оно даётся формулой

$$M'M'' = \sqrt{(x' - x'')^2 + (y' - y'')^2}. \tag{16}$$

Таким же образом, говоря о тройках чисел (x', y', z') и (x'', y'', z''), мы можем прибегнуть к пространственной координатной системе и получим расстояние между соответствующими точками M' и M'' в виде формулы

$$M'M'' = \sqrt{(x' - x'')^2 + (y' - y'')^2 + (z' - z'')^2}. \tag{17}$$

По аналогии, говоря о системах из n чисел $(x'_1, x'_2, \ldots, x'_n)$ и $(x''_1, x''_2, \ldots, x''_n)$, можно обратиться к воображаемому n-мерному

пространству и использовать расстояние между соответствующими точками M' и M'':

$$M'M'' = \sqrt{(x_1' - x_1'')^2 + (x_2' - x_2'')^2 + \ldots + (x_n' - x_n'')^2} \quad (18)$$

или в более краткой записи

$$M'M'' = \sqrt{\sum_{m=1}^{n} (x_m' - x_m'')^2}\ ^{1}). \quad (19)$$

Как мы видели в § 36, ε-окрестность точки x_0 на числовой оси Ox определяется неравенством

$$|x - x_0| < \varepsilon,$$

что даёт нам открытый промежуток

$$x_0 - \varepsilon < x < x_0 + \varepsilon.$$

На плоскости Oxy ε-окрестность точки $M_0(x_0, y_0)$ определяется неравенством

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} < \varepsilon,$$

что даёт внутренность круга с центром M_0 и радиусом ε:

$$(x - x_0)^2 + (y - y_0)^2 < \varepsilon^2.$$

В трёхмерном пространстве $Oxyz$ мы встречаемся со сферой

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < \varepsilon^2;$$

в n-мерном пространстве $Ox_1x_2 \ldots x_n$ — с «гиперсферой»

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \ldots + (x_n - x_n^0)^2 < \varepsilon^2.$$

Однако нужно заметить, что метрика, определяемая формулами вида (15) — (19), обусловливается используемыми геометрическими представлениями и не является логически необходимой. Эта метрика носит название *евклидовой*. В иных случаях целесообразным является введение иной метрики.

Поясним это на следующем примере. Пусть речь идёт о «расстояниях» между двумя точками сферы, например, земного шара. В геометрическом смысле расстояние между двумя точками сферы равно длине соединяющего их отрезка, т. е. хорды сферы; но расстояние между точками на земной поверхности измеряют по этой поверхности, именно — *по дуге большого круга, проведённого через две точки* (это — «кратчайшее» расстояние на сфере). Предположим, что две точки M' и M'' заданы их географическими координатами — широтой u и долготой v. Допустим, что мы имеем дело с картой, на которой точка $M(u, v)$ изображается точкой с декартовыми координа-

---

[1]) Следует обратить внимание на то, что при $n = 1$ формула (18) принимает вид (15).

тами u (абсцисса) и v (ордината) — см. пример 8 предыдущего параграфа. В этом случае формула $\sqrt{(u'-u'')^2+(v'-v'')^2}$ определяет расстояние между точками на карте, но не на земной поверхности; это последнее расстояние есть более сложная функция четырёх координат u', v', u'', v'', которой мы приводить не будем.

Представим себе теперь вместо шара с нанесённой на нём «географической» координатной сеткой произвольную кривую поверхность в пространстве с нанесённой на ней произвольной «криволинейной» координатной сеткой.

Расстояние между двумя точками M' и M'' на поверхности обыкновенно определяется как длина к р а т ч а й ш е й из всех линий, которые могут быть проведены на поверхности от точки M' к точке M'' [1]); это расстояние есть некоторая более или менее сложная функция от четырёх координат точек M' и M''. Оно вовсе не обязательно совпадает с расстоянием между «изображениями» соответствующих точек на плоскости, если бы мы решили криволинейные координаты на поверхности трактовать как декартовы на плоскости.

Обратимся теперь к общей теории. Говорят, что в некотором множестве («пространстве») $\mathcal{X}$ в в е д е н а м е т р и к а (мероопределение), если с каждой парой элементов X', X'' сопоставлено в качестве расстояния между ними некоторое число $\rho(X', X'')$. Это число по определению должно обладать следующими свойствами:

$$1^\circ \qquad \rho(X', X'') = \rho(X'', X'). \tag{20}$$

Это значит, что расстояние, «измеряемое от X' до X''», равно расстоянию, «измеряемому от X'' до X'»; такого рода симметрией оправдывается оборот речи «расстояние между элементами X' и X''».

$$2^\circ \qquad \left.\begin{array}{l} \rho(X', X'') > 0, \text{ если элементы } X' \text{ и } X'' \text{ различны;} \\ \rho(X', X'') = 0, \text{ если элементы } X' \text{ и } X'' \text{ совпадают.} \end{array}\right\} \tag{21}$$

Это значит, что расстояние между двумя элементами не может быть отрицательным и обращается в нуль только в том случае, если X' и X'' — один и тот же элемент.

3° Каковы бы ни были три элемента X', X'' и X''' из $\mathcal{X}$, имеет место неравенство

$$\rho(X', X'') \leqslant \rho(X', X''') + \rho(X''', X''). \tag{22}$$

Это неравенство, называемое *«неравенством треугольника»*, означает, что расстояние, «измеряемое непосредственно от X' до X''», не должно превышать суммы двух расстояний, именно, «измеряемого от X' до некоторого третьего элемента X'''» и «измеряемого от X''' до X''».

П р и м е р. В случае, если «пространство» $\mathcal{X}$ — множество всех точек на плоскости Oxy, элементы его — точки или пары чисел

---

[1]) Число таких линий бесконечно; существование кратчайшей из них требует, конечно, особого доказательства.

$M(x, y)$, то расстояние $\rho(M', M'')$ между точками $M'(x', y')$ и $M''(x'', y'')$ можно определить по формуле

$$\rho(M', M'') = \sqrt{(x'-x'')^2+(y'-y'')^2}. \tag{23}$$

Но законно было бы также взять за расстояние

$$\rho(M', M'') = |x'-x''|+|y'-y''| \tag{24}$$

или

$$\rho(M', M'') = \max\{|x'-x''|, |y'-y''|\}. \tag{25}$$

Читатель легко проверит, что в случаях (24) и (25), как и в случае (23), оправдываются свойства 1° и 2°; формальное доказательство свойства 3° несколько сложнее, но тоже вполне элементарно[1]). Читатель отдаст себе также отчёт в геометрическом смысле «расстояния» согласно толкованиям (24) и (25).

Раз в пространстве $\mathscr{X}$ введена метрика (мероопределение), тем самым определяется и окрестность каждого элемента: под *ε-окрестностью* элемента X_0 из $\mathscr{X}$ подразумевают множество всех тех элементов X из $\mathscr{X}$, для которых выполняется неравенство

$$\rho(X, X_0) < \varepsilon.$$

Пример. Смотря по тому, принята ли на плоскости метрика, определяемая формулой (23) или формулой (24), или формулой (25), ε-окрестность точки $M_0(x_0, y_0)$ определяется неравенством $(x-x_0)^2+(y-y_0)^2<\varepsilon^2$ (внутренность круга с центром M_0 и радиусом ε) или неравенством

$$\max\{|x-x_0|, |y-y_0|\} < \varepsilon$$

(внутренность квадрата с центром M_0 и стороной 2ε, причём стороны параллельны осям координат), или неравенством

$$|x-x_0|+|y-y_0| < \varepsilon$$

(внутренность квадрата с центром M_0 и вершинами $(x_0+\varepsilon, y_0)$, $(x_0, y_0+\varepsilon)$, $(x_0-\varepsilon, y_0)$ и $(x_0, y_0-\varepsilon)$).

Мы рассматривали до сих пор преимущественно такие пространства, элементами которых являются точки (числа или системы чисел). Но можно также рассматривать пространства, элементами которых являются прямые, окружности, плоскости, какие угодно кривые или поверхности и т. п. В таких «пространствах» также можно вводить метрику.

Мы остановимся лишь на нескольких отдельных примерах «функциональных» пространств, т. е. пространств, элементами которых

---

[1]) *Неравенство Минковского:*

$$\sqrt{(x'-x'')^2+(y'-y'')^2} \leqslant$$

$$\leqslant \sqrt{(x'-x''')^2+(y'-y''')^2}+\sqrt{(x'''-x'')^2+(y'''-y'')^2}.$$

являются функции, а введение метрики в таких пространствах объясним на примерах.

Станем рассматривать в качестве элементов некоторого функционального пространства функции $f(x)$ одной независимой переменной x, заданные в одном и том же промежутке $I(a \leqslant x \leqslant b)$; при этом для простоты будем допускать, что речь идёт лишь о непрерывных функциях. Итак, пусть $\mathscr{X}$ обозначает множество всех функций $f(x)$, заданных и непрерывных в промежутке I.

Как ввести метрику в пространстве $\mathscr{X}$? Нам уже пришлось встретиться (см. § 48) с «расстоянием между функциями $f(x)$ и $g(x)$», определяемым по формуле

$$\rho(f, g) = \max_I \{|f(x) - g(x)|\}; \tag{26}$$

геометрически, как мы видели, это означает, что в качестве «расстояния» берётся наибольшая длина отрезка PQ (см. рис. 87) при изменении x в пределах промежутка I.

В данной связи важно указать, что можно было бы также в качестве «расстояния» $\rho(f, g)$ взять один из интегралов

$$\int_a^b |f(x) - g(x)|\, dx \tag{27}$$

(площадь заштрихованной фигуры на рис. 87) или

$$\int_a^b [f(x) - g(x)]^2\, dx \tag{28}$$

(выбор такого «расстояния» особенно выгоден с точки зрения простоты вычислений).

Проверка свойств «расстояния» 1°—3° вполне элементарна в случае метрики (26); в случае же метрик (27) и (28) свойства 1° и 2° очевидны, а доказательство свойства 3° связано с использованием свойств определённых интегралов [1]).

§ 58. Понятие предела в метрическом пространстве

Пусть задано некоторое метрическое пространство $\mathscr{X}$ и пусть

$$\{A_n\} \equiv A_1, \ A_2, \ \ldots, \ A_n, \ \ldots \tag{29}$$

— некоторая последовательность его элементов. Если, кроме того, в пространстве $\mathscr{X}$ существует такой элемент A, что, как бы мало ни было положительное число ε, неравенству

$$\rho(X, A) < \varepsilon \tag{30}$$

[1]) Существует принципиально очень важная формальная аналогия между метриками (26), (27) и (28) в функциональном пространстве и метриками (25), (24) и (23) в точечном (двумерном) пространстве.

удовлетворяет бесконечное множество элементов последовательности $\{A_n\}$, то элемент A называется *предельным элементом* (или *точкой*) этой последовательности; если, больше того, этому неравенству удовлетворяют «почти все» элементы последовательности, т. е. все, начиная с некоторого номера ($n > N_\varepsilon$), или все, кроме конечного их числа, то говорят, что элемент A есть *предел последовательности* $\{A_n\}$. В последнем случае пишут:

$$A_n \to A \quad \text{или} \quad \lim A_n = A.$$

Последовательность элементов пространства не может иметь более одного предела.

В самом деле, из соотношений $A_n \to A$ и $A_n \to A'$ следует, что при любом $\varepsilon\,(>0)$ и при достаточно больших значениях n справедливы одновременно неравенства

$$\rho(A_n, A) < \varepsilon \quad \text{и} \quad \rho(A_n, A') < \varepsilon,$$

и отсюда (по свойствам «расстояний» 3° и 1°) вытекает:

$$\rho(A, A') \leqslant \rho(A, A_n) + \rho(A_n, A') = \\ = \rho(A_n, A) + \rho(A_n, A') < \varepsilon + \varepsilon = 2\varepsilon.$$

Но ε сколь угодно мало; следовательно,

$$\rho(A, A') = 0,$$

и потому (по свойству 2°) A' совпадает с A.

Последовательность $\{A_n\}$, имеющая предел, называется *сходящейся* (к соответствующему пределу).

Последовательность $\{A_n\}$ называется *ограниченной*, если ограничена числовая последовательность $\{\rho(A_n, A')\}$, где A' — некоторый элемент $\mathscr{X}$. Понятие «ограниченная последовательность» не зависит от выбора элемента A', так как, если ограничена последовательность $\{\rho(A_n, A')\}$, то ограничена и последовательность $\{\rho(A_n, A'')\}$ (и обратно).

Действительно, пусть

$$\rho(A_n, A') < M \qquad (n = 1, 2, 3, \ldots).$$

Полагая

$$\rho(A', A'') = m,$$

будем иметь (по свойству 3°):

$$\rho(A_n, A'') \leqslant \rho(A_n, A') + \rho(A', A'') < M + m.$$

Возникает вопрос: верно ли, что всякая ограниченная последовательность $\{A_n\}$ непременно должна иметь предельный элемент? Всегда ли возможно, другими словами, обобщение теоремы Больцано-Вейерштрасса (§ 36) на произвольные метрические пространства?

Ответ на этот вопрос — отрицательный. Теорема Больцано-Вейерштрасса не распространяется, например, на пространство непрерывных функций, с «равномерной» метрикой, определяемой по формуле (26). В самом деле, рассмотрим хотя бы последовательность функций

$$\{\sin nx\}. \tag{31}$$

Эта последовательность — ограниченная; взяв в качестве «начального элемента» A' хотя бы функцию $f \equiv 0$, мы получаем:

$$\rho(\sin nx,\ 0) = \max|\sin nx| \leqslant 1.$$

Но в основном промежутке $a \leqslant x \leqslant b$ эта последовательность, как легко понять, не имеет ни одного предельного элемента.

Говорят, что некоторое множество A элементов метрического пространства $\mathscr{X}$ называется *компактным*, если всякая последовательность $\{A_n\}$, составленная из элементов A, имеет предельный элемент.

Если множество $\mathscr{A}$ элементов метрического пространства компактно, то оно непременно ограничено.

В самом деле, в противном случае можно из $\mathscr{A}$ извлечь такую последовательность элементов $\{A_n\}$, что

$$\rho(A_n,\ A') \to \infty \tag{32}$$

(где A' — некоторый элемент $\mathscr{A}$). С другой стороны, по свойству компактности $\mathscr{A}$ последовательность $\{A_n\}$ имеет предельный элемент, например A; следовательно, как бы мало ни было ε, бесконечное множество элементов последовательности $\{A_n\}$ удовлетворяют неравенству

$$\rho(A_n,\ A) < \varepsilon.$$

Для этих элементов мы получили бы, пользуясь неравенством треугольника,

$$\rho(A_n,\ A') \leqslant \rho(A_n,\ A) + \rho(A,\ A') < \varepsilon + \rho(A,\ A') \equiv M,$$

а это противоречит соотношению (32).

Из вышеприведённого примера видно, что в пространстве непрерывных функций с «равномерной» метрикой совокупность элементов ограниченной последовательности может не обладать свойством компактности.

Но мы видели (§ 36), что таким свойством обладает совокупность элементов всякой ограниченной последовательности в точечном одномерном пространстве. Легко убедиться, что это справедливо и для всякого конечномерного точечного пространства.

Проведём доказательство последнего утверждения хотя бы для случая плоскости: обобщение на n-мерное пространство совершенно очевидно. Пусть дана ограниченная последовательность точек

$$\{P_n(x_n,\ y_n)\}.$$

Принимая в качестве «начального элемента» A' начало координат O, мы видим, что по свойству ограниченности

$$\rho(P_n,\ 0) = \sqrt{x_n^2 + y_n^2} < M \qquad (n = 1,\ 2,\ 3, \ldots);$$

но тогда одновременно

$$|x_n| < M \tag{33}$$

и

$$(n = 1,\ 2,\ 3, \ldots).$$

$$|y_n| < M \tag{34}$$

Так как числовая последовательность $\{x_n\}$ согласно (33) ограничена, то она имеет предельную точку, например ξ. Это значит, что в промежутке

$$\xi - \frac{\varepsilon}{\sqrt{2}} < x < \xi + \frac{\varepsilon}{\sqrt{2}} \tag{35}$$

имеется бесконечное множество чисел x_n, например

$$x_{p_1}, x_{p_2}, \ldots, x_{p_n}, \ldots$$

Рассмотрим теперь числовую последовательность

$$y_{p_1}, y_{p_2}, \ldots, y_{p_n}, \ldots$$

Согласно (34) она — ограниченная и, следовательно, имеет предельную точку, скажем η. Это значит, что в промежутке

$$\eta - \frac{\varepsilon}{\sqrt{2}} < y < \eta + \frac{\varepsilon}{\sqrt{2}} \tag{36}$$

имеется бесконечное множество чисел y_{p_n}. Отсюда ясно, что в области, определяемой совокупностью неравенств (35) и (36), имеется бесконечное множество точек P_n. Но эта область заключена целиком в круговой области

$$(x-\xi)^2 + (y-\eta)^2 < \varepsilon^2,$$

являющейся ε-окрестностью точки (ξ, η). Так как ε произвольно мало, то эта точка есть предельный элемент нашей последовательности точек.

В компактном метрическом пространстве $\mathcal{X}$ предел последовательности элементов может быть определён как единственная предельная точка.

В самом деле, пусть A есть единственная предельная точка последовательности $\{A_n\}$; убедимся, что она есть предел этой последовательности, т. е. что при некотором $\varepsilon > 0$ лишь конечное число элементов A_n удовлетворяют неравенству

$$\rho(A_n, A) \geqslant \varepsilon. \tag{37}$$

Если бы число таких элементов было бесконечным, то по свойству компактности $\mathcal{X}$ всякая образованная из них последовательность имела бы хоть одну предельную точку B. В таком случае бесконечное множество элементов нашей последовательности удовлетворяло бы одновременно неравенству

$$\rho(A_n, B) < \frac{\varepsilon}{2}$$

и вместе с тем неравенству (37); и так как по свойству 3°

$$\rho(A_n, A) \leqslant \rho(A, B) + \rho(B, A_n),$$

то отсюда следовало бы заключение

$$\rho(A, B) \geqslant \rho(A_n, A) - \rho(A_n, B) > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Но тогда ясно, что элементы A и B различны, так что последовательность $\{A_n\}$ вопреки предположению имела бы по крайней мере две предельные точки.

Обратно, если последовательность $\{A_n\}$ имеет предел A, то она не может иметь предельной точки B, отличной от A. Действительно, в противном случае, взяв ε по условию

$$0 < \varepsilon < \frac{1}{2}\rho(A, B), \tag{38}$$

мы имели бы неравенства $\rho(A_n, A) < \varepsilon$ (для достаточно больших n) и $\rho(A_n, B) < \varepsilon$ (для сколь угодно больших n), откуда следовало бы

$$\rho(A, B) \leqslant \rho(A, A_n) + \rho(A_n, B) < 2\varepsilon,$$

что противоречит неравенству (38).

Напротив, если пространство $\mathscr{X}$ не обладает свойством компактности, единственная предельная точка последовательности может не быть пределом. Примером может служить последовательность в пространстве непрерывных функций с равномерной сходимостью

$$\sin x,\ 0,\ \sin 2x,\ 0,\ \sin 3x,\ 0,\ \ldots,\ \sin nx,\ 0,\ \ldots; \tag{39}$$

она имеет единственный предельный элемент $f \equiv 0$, который, однако, не является пределом последовательности.

Заметим, что по поводу неограниченных последовательностей нужно сказать, что предела в собственном смысле они иметь не могут, однако в этом случае возможны различные обобщения понятия предела («предел в несобственном смысле», см. § 39).

§ 59. Топологические пространства [1])

Если в пространстве введена метрика с соблюдением условий 1°—3° § 57, то тем самым устанавливается, как мы только что убедились, понятие предела последовательности элементов этого пространства.

Однако понятие предела последовательности может быть введено и независимо от метрики. Так, например, понятие простого (т. е. без требования равномерности) предела последовательности функций $\{f_n(x)\}$, заданных в данном промежутке $I(a \leqslant x \leqslant b)$, не связано с предварительным введением метрики (см. § 45).

Говорят, что некоторое пространство $\mathscr{X}$ — *топологическое*, если тем или иным способом в нём введено понятие предела последовательности элементов. При этом точный смысл слов «введено

[1]) Читатель должен быть предупреждён, что термин «топологическое пространство» в данной статье употребляется в несколько более широком смысле, чем в современной топологии.

понятие предела» заключается в следующем: устанавливается правило, согласно которому с каждой данной последовательностью $\{A_n\}$ элементов $\mathcal{X}$ или не сопоставляется никакого элемента из $\mathcal{X}$ или сопоставляется некоторый, только один, элемент A из $\mathcal{X}$, называемый «пределом» последовательности $\{A_n\}$. Упомянутое правило должно удовлетворять единственному требованию: *если A есть предел последовательности элементов $\{A_n\}$, то A есть вместе с тем предел любой подпоследовательности* («частной последовательности») *элементов $\{A_{p_n}\}$*. Короче: *если* $A_n \to A$, *то и* $A_{p_n} \to A$. Легко понять, что множество всех функций $f(x)$, заданных в промежутке I (или некоторая часть этого множества), представляет собой топологическое пространство, если в качестве правила, устанавливающего понятие предела, ввести простую сходимость. Вместе с тем топологическим является также любое метрическое пространство.

После того, как пространство $\mathcal{X}$ «топологизировано» (т. е. в нём введено понятие предела), может быть введено и понятие предельного элемента последовательности. Именно, говорят, что элемент A из $\mathcal{X}$ есть *предельный элемент* последовательности $\{A_n\}$ элементов из $\mathcal{X}$, если существует подпоследовательность $\{A_{p_n}\}$, для которой A есть предел: $A_{p_n} \to A$. Очевидно, *предел последовательности $\{A_n\}$* (если он существует) *есть вместе с тем её предельный элемент, и притом — единственный.*

В самом деле, если бы существовал ещё иной предельный элемент, например A', то для некоторой подпоследовательности $\{A_{p_n}\}$ мы имели бы $A_{p_n} \to A'$; и так как по определению предела для всякой подпоследовательности $A_{p_n} \to A$, то отсюда следовало бы, что вопреки предположению A' и A совпадают.

Введённое выше определение предельного элемента точно соответствует свойствам предельного элемента в метрическом пространстве.

Действительно, если A есть предельный элемент последовательности $\{A_n\}$ в метрическом пространстве $\mathcal{X}$, то, как бы мало ни было ε, в окрестности $\rho(X, A) < \varepsilon$ существует бесконечное множество элементов A_n. Выберем такую последовательность положительных чисел $\{\varepsilon_n\}$, что $\varepsilon_n \to 0$; возьмём какой-нибудь элемент A_{p_1} из окрестности $\rho(X, A) < \varepsilon_1$, затем какой-нибудь элемент A_{p_2} из окрестности $\rho(X, A) < \varepsilon_2$ (с ограничением $p_2 > p_1$); затем, элемент A_{p_3} из окрестности $\rho(X, A) < \varepsilon_3$ (с ограничением $p_3 > p_2$) и т. д. Тогда последовательность $\{A_{p_n}\}$ обладает свойством

$$\rho(A_{p_n}, A) < \varepsilon_n,$$

так что

$$\rho(A_{p_n}, A) \to 0,$$

т. е.

$$A_{p_n} \to A.$$

Обратно, если существует такая подпоследовательность $\{A_{p_n}\}$, что $A_{p_n} \to A$, то A есть предельный элемент последовательности $\{A_n\}$. В самом деле, соотношение $A_{p_n} \to A$ означает, что при достаточно больших n

$$\rho(A_{p_n}, A) < \varepsilon,$$

так что в ε-окрестности точки A содержится бесконечное множество элементов последовательности $\{A_n\}$.

Понятие компактного множества переносится на общий случай топологического пространства без изменений.

Следует отметить, что в общем топологическом пространстве единственный предельный элемент последовательности не обязательно является её пределом — даже в том случае, если все элементы последовательности принадлежат некоторому компактному множеству $\mathscr{A}$ [1]).

§ 60. Алгебра множеств. Производное множество. Замкнутость и связность

Предполагая рассматривать одновременно различные множества, составленные из элементов некоторого пространства $\mathscr{X}$, мы должны условиться в некоторых обозначениях, которыми придётся пользоваться ради краткости речи и ясности понимания. Приведём их в табличной форме.

Если $\mathscr{A}$, $\mathscr{B}$ и т. д. — множества элементов из $\mathscr{X}$, то запись	Означает:
$\mathscr{A} \subseteq \mathscr{B}$ или $\mathscr{B} \supseteq \mathscr{A}$	Всякий элемент множества $\mathscr{A}$ есть вместе с тем элемент множества $\mathscr{B}$ (читается: «$\mathscr{A}$ входит в $\mathscr{B}$» или «$\mathscr{A}$ есть часть $\mathscr{B}$»).
$\mathscr{A} \subset \mathscr{B}$ или $\mathscr{B} \supset \mathscr{A}$	То же самое, но при этом существует хоть один элемент множества $\mathscr{B}$, не являющийся элементом множества $\mathscr{A}$ (читается: «$\mathscr{A}$ есть истинная часть $\mathscr{B}$»).
$\mathscr{A} = \mathscr{B}$	$\mathscr{A} \subseteq \mathscr{B}$ и вместе с тем $\mathscr{A} \supseteq \mathscr{B}$, т. е. множества $\mathscr{A}$ и $\mathscr{B}$ состоят из одних и тех же элементов.

[1]) В самом деле, пусть правило, определяющее предел, таково: последовательность $\{A_n\}$ имеет предел A, если все элементы последовательности совпадают с A. Множество, состоящее из двух различных элементов, например P и Q, очевидно, компактно; и тем не менее последовательность, у которой первый элемент есть P, а все остальные — Q, имеет единственный предельный элемент Q, но не имеет предела.

Если $\mathscr{A}$, $\mathscr{B}$ и т. д. — множества элементов из $\mathscr{X}$, то запись	Означает:
$\mathscr{A}+\mathscr{B}$	Множество, составленное из всех элементов, принадлежащих или множеству $\mathscr{A}$ или множеству $\mathscr{B}$ («сумма» или «объединение» множеств $\mathscr{A}$ и $\mathscr{B}$).
$\mathscr{A}\mathscr{B}$	Множество, составленное из всех элементов, принадлежащих и множеству $\mathscr{A}$ и множеству $\mathscr{B}$ («произведение» или «пересечение» множеств $\mathscr{A}$ и $\mathscr{B}$).
0	«Пустое» множество (не содержащее ни одного элемента).

Скобки, как в обыкновенной алгебре, указывают порядок действий.

Справедливы утверждения:

1) если $\mathscr{A}\subseteq\mathscr{B}$ и $\mathscr{B}\subseteq\mathscr{C}$, то $\mathscr{A}\subseteq\mathscr{C}$;

2) если $\mathscr{A}=\mathscr{B}$ и $\mathscr{B}=\mathscr{C}$, то $\mathscr{A}=\mathscr{C}$;

3) $\mathscr{A}+\mathscr{B}=\mathscr{B}+\mathscr{A}$;

4) $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}$;

5) $\mathscr{A}+(\mathscr{B}+\mathscr{C})=(\mathscr{A}+\mathscr{B})+\mathscr{C}$;

6) $\mathscr{A}\ (\mathscr{B}\mathscr{C})=(\mathscr{A}\mathscr{B})\ \mathscr{C}$;

7) $\mathscr{A}\ (\mathscr{B}+\mathscr{C})=\mathscr{A}\mathscr{B}+\mathscr{A}\mathscr{C}$;

8) $\mathscr{A}+\mathscr{B}\mathscr{C}=(\mathscr{A}+\mathscr{B})\ (\mathscr{A}+\mathscr{C})$;

9) $\mathscr{A}\supseteq\mathscr{A}\mathscr{B}$;

10) $\mathscr{A}\subseteq\mathscr{A}+\mathscr{B}$;

11) если $\mathscr{A}\subseteq\mathscr{B}$, то $\mathscr{A}\mathscr{B}=\mathscr{A}$, и обратно;

12) если $\mathscr{A}\subseteq\mathscr{B}$, то $\mathscr{A}+\mathscr{B}=\mathscr{B}$, и обратно;

13) $\mathscr{A}+\mathscr{A}=\mathscr{A}\mathscr{A}=\mathscr{A}$;

14) если $\mathscr{A}+\mathscr{B}=0$, то $\mathscr{A}=\mathscr{B}=0$, и обратно;

15) если $\mathscr{A}\mathscr{B}=\mathscr{X}$, то $\mathscr{A}=\mathscr{B}=\mathscr{X}$, и обратно.

Предполагая далее, что пространство $\mathscr{X}$ — топологическое, введём понятие предельного элемента множества $\mathscr{A}$ и понятие производного множества от множества $\mathscr{A}$. Элемент A пространства $\mathscr{X}$ (принадлежащий или не принадлежащий множеству $\mathscr{A}$) называется *предельным элементом множества* $\mathscr{A}$, если можно указать

последовательность $\{A_n\}$ различных между собой элементов $\mathscr{A}$, имеющую предел A:

$$\mathscr{A}_n \to \mathscr{A}\ ^{1}).$$

Множество всех предельных элементов данного множества $\mathscr{A}$ называется *производным множеством* множества $\mathscr{A}$ и обозначается через $\mathscr{A}'$.

Приведём несколько примеров.

1. При обычной метрике, если $\mathscr{A}$ есть открытый промежуток $a<x<b$, то $\mathscr{A}'$ — замкнутый промежуток $a\leqslant x\leqslant b$.

2. Тот же результат, если $\mathscr{A}$ — множество конечных десятичных дробей в промежутке $a<x<b$.

3. Если $\mathscr{A}$ — множество точек (x, y), удовлетворяющих условию $\begin{cases}0\leqslant x\leqslant 1\\ 0\leqslant y\leqslant 1\end{cases}$ (внутренность квадрата), то $\mathscr{A}'$ — множество точек, удовлетворяющих условию $\begin{cases}0\leqslant x\leqslant 1,\\ 0\leqslant y\leqslant 1\end{cases}$ (квадрат вместе с контуром).

4. Если $\mathscr{A}$ — множество точек кривой $y=\sin\frac{1}{x}$ $(0<x\leqslant 1)$ (см. § 41, пример 4), то $\mathscr{A}'$ — множество тех же точек, с добавлением отрезка $\begin{cases}x=0,\\ -1\leqslant y\leqslant +1.\end{cases}$

5. Если $\mathscr{A}$ — множество дробей вида $\frac{1}{n}$ (n — целое положительное), то $\mathscr{A}'$ состоит из одного числа 0.

6. Если $\mathscr{A}$ — множество дробей вида $\frac{1}{p}+\frac{1}{q}$ (p и q — целые положительные), то $\mathscr{A}'$ — множество, составленное из дробей вида $\frac{1}{p}$ и из числа 0.

7. Если $\mathscr{A}$ — множество точек окружностей $x^2+y^2=\left(1-\frac{1}{n}\right)^2$ $(n=2, 3, 4, \ldots)$, то $\mathscr{A}'$ — то же множество, с добавлением точек окружности $x^2+y^2=1$.

8. Если в метрике равномерной сходимости $\mathscr{A}$ — множество всех многочленов (рассматриваемых в данном промежутке $a\leqslant x\leqslant b$), то $\mathscr{A}'$ — множество всех непрерывных функций, заданных в том же промежутке (теорема Вейерштрасса, см. § 49).

---

[1]) Таким образом, не одно и то же — «предельный элемент последовательности» и «предельный элемент множества элементов последовательности». Например, последовательность чисел 0, 1, 0, 1, 0, 1, ... имеет предельные элементы 0 и 1, но множество, составленное из элементов этой последовательности, содержит лишь два элемента и потому (как и всякое конечное множество) не имеет предельных элементов.

Для последовательности «без повторяющихся элементов» понятие предельного элемента последовательности и предельного элемента множества элементов последовательности совпадают.

9. Если в метрике простой сходимости $\mathscr{A}$ — множество функций вида $f_n(x) = \frac{1}{1+n^2x^2}$ $(n = 1, 2, 3, \ldots)$, то $\mathscr{A}'$ состоит из единственной функции

$$f(x) = \begin{cases} 1 \text{ при } x = 0, \\ 0 \text{ при } x \neq 0. \end{cases}$$

В метрике равномерной сходимости множество $\mathscr{A}'$ пусто.

Множество $\mathscr{A}$ в топологическом пространстве $\mathscr{X}$ называется *замкнутым*, если

$$\mathscr{A}' \subseteq \mathscr{A}.$$

Множество

$$\overline{\mathscr{A}} = \mathscr{A} + \mathscr{A}'$$

иногда называют *замыканием* множества $\mathscr{A}$. Замыкание замкнутого множества, очевидно, совпадает с самим множеством; и обратно, если замыкание множества совпадает с самим множеством, то это множество — замкнутое.

Примерами замкнутых множеств могут служить:

а) Множество точек, принадлежащих замкнутому отрезку $a \leqslant x \leqslant b$.

б) Множество непрерывных функций $f(x)$, заданных на отрезке $a \leqslant x \leqslant b$ при равномерной метрике (см. § 48).

Заметим, что из $\mathscr{A} \subseteq \mathscr{B}$ следует $\mathscr{A}' \subseteq \mathscr{B}'$ (очевидно).

Теорема. *В метрическом пространстве всякое производное множество замкнуто.*

В виде формулы это утверждение записывается так:

$$\mathscr{A}'' \subseteq \mathscr{A}'.$$

Доказательство. Пусть A'' есть некоторый элемент множества $\mathscr{A}''$. Нужно показать, что, как бы мало ни было ε, существует такой элемент A из $\mathscr{A}$, что

$$\rho(A, A'') < \varepsilon.$$

Так как A'' есть предельный элемент множества $\mathscr{A}'$, то существует такой элемент A' из $\mathscr{A}'$, что

$$\rho(A', A'') < \frac{\varepsilon}{2}.$$

С другой стороны, так как A' есть предельный элемент множества $\mathscr{A}$, то существует такой элемент A из $\mathscr{A}$, что

$$\rho(A, A') < \frac{\varepsilon}{2}.$$

В таком случае, по свойству треугольника,

$$\rho(A, A'') \leqslant \rho(A, A') + \rho(A', A'') < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

что и требовалось доказать.

Эта теорема, как легко понять, обобщает на произвольные метрические пространства теорему, сформулированную на стр. 163.

В произвольном топологическом пространстве теорема неверна.

Множество $\mathscr{A}$ в топологическом пространстве $\mathscr{X}$ называется *связным*, если нельзя указать двух таких множеств $\mathscr{M}$ и $\mathscr{N}$ в том же пространстве, чтобы выполнялись условия:

$$\left.\begin{array}{l} 1^\circ\ \mathscr{A} = \mathscr{M} + \mathscr{N}, \\ 2^\circ\ \mathscr{M}\mathscr{N} + \mathscr{M}\mathscr{N}' + \mathscr{M}'\mathscr{N} = 0, \\ 3^\circ\ \mathscr{M} \neq 0, \quad \mathscr{N} \neq 0. \end{array}\right\} \tag{40}$$

Примеры (относящиеся к случаю обыкновенной евклидовой метрики).

10. Множество, состоящее из двух чисел a и b $(a \neq b)$, не является связным: множество $\mathscr{M}$ можно взять состоящим из одного элемента a, множество $\mathscr{N}$ — состоящим из одного элемента b.

11. Аналогично в случае множества, составленного из точек двух кругов

$$(x+c)^2 + y^2 < R^2 \quad \text{и} \quad (x-c)^2 + y^2 < R^2,$$

при условии $c > R > 0$ и даже при $c = R > 0$.

12. Если числовое множество $\mathscr{A}$ содержит элементы a и c, но не содержит b, причём $a < b < c$, то оно — не связное. В качестве $\mathscr{M}$ можно взять совокупность точек множества x, для которых $x < b$; в качестве $\mathscr{N}$ — совокупность тех точек x, для которых $x > b$.

13. Числовое множество (замкнутый промежуток) $a \leqslant x \leqslant b$ связно.

Это утверждение могло бы показаться очевидным, однако формальное доказательство его не так просто. Наметим ход доказательства «от противного», предоставляя читателю воспроизвести его во всех деталях.

Принимая промежуток $a \leqslant x \leqslant b$ за $\mathscr{A}$, из условий 1° и 3° (40) заключаем о существовании точки m_1, принадлежащей множеству $\mathscr{A}\mathscr{M}$, и точки n_1, принадлежащей множеству $\mathscr{A}\mathscr{N}$. Из условия 2° следует: $\mathscr{M}\mathscr{N} = 0$. Поэтому $m_1 \neq n_1$; пусть, например, $m_1 < n_1$. Точка $\frac{m_1 + n_1}{2}$ принадлежит $\mathscr{A}$; значит, принадлежит или $\mathscr{M}$ или $\mathscr{N}$. В первом случае положим $m_2 = \frac{m_1 + n_1}{2}$, $n_2 = n_1$, во втором $m_2 = m_1$, $n_2 = \frac{m_1 + n_1}{2}$; затем будем так же рассуждать по поводу m_2 и n_2. Продолжая эту процедуру, убедимся, что общий предел ξ монотонных последовательностей чисел m и чисел n будет принадлежать множеству $\mathscr{M}'\mathscr{N}'$. Принимая во внимание, что само ξ принадлежит или $\mathscr{M}$ или $\mathscr{N}$, получим противоречие с равенствами

$$\mathscr{M}\mathscr{N}' = \mathscr{M}'\mathscr{N} = 0.$$

14. Сказанное о замкнутом промежутке переносится на любой промежуток.

Обратно, всякое связное множество точек на числовой оси является промежутком (конечным или бесконечным, замкнутым или незамкнутым, или «полузамкнутым», т. е. замкнутым лишь с одного конца; см. выше пример 12).

Множество $\mathscr{A}$ в пространстве $\mathscr{X}$, связное и замкнутое в этом пространстве, называется *континуумом*.

Итак, если под пространством $\mathscr{X}$ понимать «одномерное евклидово пространство», т. е. числовую прямую с обычной метрикой, то в этом пространстве континуумы исчерпываются замкнутыми промежутками.

В «двумерном евклидовом пространстве», т. е. на плоскости с обычной метрикой, имеются более разнообразные типы континуумов: сюда относится, например, всякий квадрат со включением сторон и вершин или ещё всякий круг со включением контура (окружности).

§ 61. Непрерывные отображения и их свойства

Пусть топологическое пространство $\mathscr{X}$ отображено на топологическое пространство $\mathscr{Y}$. Условимся через $Y=f(X)$ обозначать тот элемент $\mathscr{Y}$, который соответствует элементу X пространства $\mathscr{X}$.

Если $\mathscr{A}$ есть некоторое множество элементов $\mathscr{X}$, то через $f(\mathscr{A})$ условимся обозначать совокупность элементов множества $\mathscr{Y}$, соответствующих элементам множества $\mathscr{A}$.

Отображение

$$Y=f(X)$$

называется *непрерывным* относительно элемента X, если из соотношения

$$X_n \to X$$

вытекает соотношение

$$f(X_n) \to f(X);$$

оно называется *непрерывным на множестве* $\mathscr{A}$, если непрерывно относительно любого элемента множества $\mathscr{A}$.

Установим несколько свойств непрерывных отображений.

Теорема I. *Если множество* $\mathscr{A}$ $(\subseteq \mathscr{X})$ *связно и отображение* $Y=f(X)$ *непрерывно на множестве* $\mathscr{A}$, *то множество* $\mathscr{B}=f(\mathscr{A})$ $(\subseteq \mathscr{Y})$ *также связно.*

Будем доказывать от противного. Пусть $\mathscr{B}$ не связно. Тогда существуют такие множества $\mathscr{M}$ и $\mathscr{N}$, что

$$1)\ \mathscr{B}=\mathscr{M}+\mathscr{N},$$

$$2)\ \mathscr{M}\mathscr{N}+\mathscr{M}\mathscr{N}'+\mathscr{M}'\mathscr{N}=0,$$

$$3)\ \mathscr{M}\neq 0,\ \mathscr{N}\neq 0.$$

Так как $\mathscr{M}\mathscr{N} = 0$, то ни один элемент A множества $\mathscr{A}$ не отображается на некоторый элемент множества $\mathscr{M}$ и на некоторый элемент множества $\mathscr{N}$ одновременно. Обозначим через $\mathscr{K}$ множество тех элементов A, которые отображаются на $\mathscr{M}$; через $\mathscr{L}$ — множество тех элементов A, которые отображаются на $\mathscr{N}$.

Очевидно, $\mathscr{K} \neq 0$ и $\mathscr{L} \neq 0$ (так как $\mathscr{M} \neq 0$ и $\mathscr{N} \neq 0$), и притом $\mathscr{K}\mathscr{L} \neq 0$. Докажем, что $\mathscr{K}\mathscr{L}' = 0$. Если бы было $\mathscr{K}\mathscr{L}' \neq 0$, то существовал бы элемент A множества $\mathscr{K}$, который был бы предельным элементом множества $\mathscr{L}$, и тогда из множества $\mathscr{L}$ можно было бы выделить такую последовательность элементов $\{A_n\}$, которая имела бы пределом элемент A множества $\mathscr{K}$:

$$A_n \to A.$$

В таком случае вследствие непрерывности отображения мы получили бы

$$f(A_n) \to f(A).$$

Это значит, что существовала бы последовательность $\{f(A_n)\}$ элементов множества $\mathscr{N}$, которая имела бы пределом элемент множества $\mathscr{M}$; но тогда мы имели бы $\mathscr{M}\mathscr{N}' \neq 0$, что противоречит условию 2°. Итак, $\mathscr{K}\mathscr{L}' = 0$; и точно так же $\mathscr{K}'\mathscr{L} = 0$. В итоге мы получаем:

1) $\mathscr{A} = \mathscr{K} + \mathscr{L}$,

2) $\mathscr{K}\mathscr{L} + \mathscr{K}\mathscr{L}' + \mathscr{K}'\mathscr{L} = 0$,

3) $\mathscr{K} \neq 0$, $\mathscr{L} \neq 0$,

т. е. выходит, что множество $\mathscr{A}$ не связно, вопреки сделанному предположению. Теорема I доказана.

Теорема II. *Если множество* $\mathscr{A}$ $(\subseteq \mathscr{X})$ *компактно и замкнуто и отображение* $Y = f(X)$ *непрерывно на* $\mathscr{A}$, *то множество* $\mathscr{B} = f(\mathscr{A})$ $(\subseteq \mathscr{Y})$ *компактно.*

Теорема III. *При тех же предположениях множество* $\mathscr{B} = f(\mathscr{A})$ $(\subseteq \mathscr{Y})$ *замкнуто.*

Пусть $\{B_n\}$ — некоторая последовательность элементов $\mathscr{B}$. Обозначим через A_n какой-нибудь элемент $\mathscr{A}$ из числа отображаемых на B_n; таким образом,

$$B_n = f(A_n) \qquad (n = 1, 2, 3, \ldots).$$

Рассмотрим последовательность $\{A_n\}$ элементов $\mathscr{A}$. Так как множество $\mathscr{A}$ компактно, то можно выделить подпоследовательность $\{A_{p_n}\}$, имеющую предел в $\mathscr{X}$; обозначим его через A. Так как $\mathscr{A}$ замкнуто, то A принадлежит множеству $\mathscr{A}$. Вследствие того, что отображение непрерывно на множестве $\mathscr{A}$, из соотношения

$$A_{p_n} \to A$$

следует соотношение

$$f(A_{p_n}) \to f(A). \tag{41}$$

Мы видим, что последовательность $\{f(A_{p_n})\}$, т. е. $\{B_{p_n}\}$, — сходящаяся: всякая последовательность $\{B_n\}$ элементов $\mathscr{B}$ имеет предельный элемент. Значит, множество $\mathscr{B}$ — компактное, и теорема II доказана.

Чтобы доказать теорему III, возобновим то же рассуждение при дополнительном предположении

$$B_n \to B, \tag{42}$$

где B — некоторый элемент пространства $\mathscr{Y}$. Соотношению (41) можно придать вид

$$B_{p_n} \to f(A), \tag{43}$$

и так как предел последовательности $\{B_{p_n}\}$ по определению предела не может быть отличен от предела последовательности $\{B_n\}$, то мы получаем из (42) и (43):

$$B = f(A).$$

Но A есть элемент $\mathscr{A}$, и раз он отображается на B, то, следовательно, B принадлежит множеству $\mathscr{B}$, что и требовалось доказать.

Как весьма частный случай, можно допустить, что топологические пространства $\mathscr{X}$ и $\mathscr{Y}$ — числовые множества, составленные из всех действительных чисел, с обычной метрикой; множество $\mathscr{A}$ — промежуток, какой угодно (применительно к теореме I) или замкнутый (применительно к теоремам II и III).

Тогда, понимая под $f(x)$ обыкновенную функцию одного действительного переменного, заданную и непрерывную в рассматриваемом промежутке, и применяя теоремы I, II и III, мы приходим к заключению (соответственно), что множество значений, принимаемых функцией $f(x)$ в этом промежутке:

1) связно, т. е. представляет собой некоторый промежуток (см. § 60, пример 14);

кроме того, при дополнительном допущении замкнутости рассматриваемого промежутка, оно:

2) компактно, т. е. ограничено (см. § 36 и § 57);

3) замкнуто, и потому существует наибольшее и наименьшее среди принимаемых значений.

Это как раз теоремы I, II и III § 47. Но теперь им придана чрезвычайная общность.

Теорему IV § 47 — о равномерной непрерывности также можно обобщить; но она по существу — метрического содержания, и потому обобщение производится на метрические, а не на произвольные топологические пространства.

Теорема IV. *Пусть компактное и замкнутое множество $\mathscr{A}$ в метрическом пространстве $\mathscr{X}$ посредством соотношения $Y = f(X)$ непрерывно отображено на некоторое множество $\mathscr{B} = f(\mathscr{A})$ в метрическом же пространстве $\mathscr{Y}$. Тогда, как бы мало ни*

было $\varepsilon\ (>0)$, *можно указать такое число* δ, *что из неравенства*

$$\rho(X', X'') < \delta \tag{44}$$

(где X' и X'' — произвольные элементы $\mathscr{A}$) *следует неравенство*

$$\rho(f(X'), f(X'')) < \varepsilon. \tag{45}$$

Доказательство аналогично приведённому на стр. 220—221. Пусть не со всяким ε можно сопоставить число δ, обладающее требуемым свойством. Значит, существует такое ε^* (>0), что к каждому числу из последовательности $\{\delta_n\}$ (где $\delta_n > 0$, $\delta_n \to 0$) можно подобрать такие пары элементов X_n', X_n'' из $\mathscr{A}$, что

$$\rho(X_n', X_n'') < \delta_n$$

и, однако,

$$\rho(f(X_n'), f(X_n'')) \geqslant \varepsilon^*. \tag{46}$$

Так как $\mathscr{A}$ компактно, то из последовательности $\{X_n'\}$ можно выделить такую сходящуюся подпоследовательность $\{X_{p_n}'\}$, что

$$X_{p_n}' \to \Xi, \tag{47}$$

причём вследствие замкнутости $\mathscr{A}$ элемент Ξ принадлежит множеству $\mathscr{A}$. Тогда и

$$X_{p_n}'' \to \Xi; \tag{48}$$

в самом деле,

$$\rho(X_{p_n}'', \Xi) \leqslant \rho(X_{p_n}'', X_{p_n}') + \rho(X_{p_n}', \Xi),$$

и так как оба слагаемых справа стремятся к нулю, то стремится к нулю и левая часть.

Из соотношений (47) и (48), далее, следует по непрерывности:

$$f(X_{p_n}') \to f(\Xi), \quad f(X_{p_n}'') \to f(\Xi). \tag{49}$$

Неравенство (46), в частности, нам даёт:

$$\rho(f(X_{p_n}'), f(X_{p_n}'')) \geqslant \varepsilon^*,$$

и теперь противоречие налицо, так как из соотношений (49) легко следует:

$$\rho(f(X_{p_n}'), f(X_{p_n}'')) \to 0.$$

§ 62. Гомеоморфные отображения

Предположим, что два множества $\mathscr{X}$ и $\mathscr{Y}$ отображены одно на другое взаимно однозначно и взаимно непрерывно. Это значит:

1° С каждым элементом X из $\mathscr{X}$ сопоставлен один и только один элемент $Y = f(X)$ из $\mathscr{Y}$ и, обратно, с каждым элементом Y

из $\mathcal{Y}$ сопоставлен один и только один элемент $X=g(Y)$ из $\mathcal{X}$, именно тот, который отображается на Y (см. § 54).

2° Если последовательность $\{X_n\}$ элементов из $\mathcal{X}$ связана с элементом X, также из $\mathcal{X}$, соотношением

$$X_n \to X, \tag{50}$$

то последовательность элементов $\{Y_n\}$, где $Y_n=f(X_n)$, связана с элементом Y, где $Y=f(X)$ (очевидно, также принадлежащим $\mathcal{Y}$), соотношением

$$Y_n \to Y, \tag{51}$$

т. е.

$$f(X_n) \to f(X),$$

и обратно: из соотношения вида (51) следует соотношение вида (50), где $X_n=g(Y_n)$, $X=g(Y)$.

Такого рода отображение множества $\mathcal{X}$ на множество $\mathcal{Y}$ (и также множества $\mathcal{Y}$ на множество $\mathcal{X}$) называется *гомеоморфным отображением*. Будем говорить кратко, что множества $\mathcal{X}$ и $\mathcal{Y}$ гомеоморфны между собой, если между ними можно установить гомеоморфное отображение.

В гомеоморфном отображении элементы X (из $\mathcal{X}$) и Y (из $\mathcal{Y}$), взаимно друг с другом сопоставляемые, называются *соответственными* (*гомологическими*).

Простейший пример гомеоморфного отображения даётся подобием двух фигур F и F'; именно, гомеоморфное отображение в этом случае таково, что, каковы бы ни были точки M и N фигуры F, между ними и им соответственными точками M' и N' фигуры F' существует зависимость

$$\frac{M'N'}{MN}=\lambda,$$

где $\lambda\,(>0)$ — постоянный коэффициент подобия. Допуская для простоты, что фигуры F и F' не только подобны, но и подобно расположены относительно начала координат O, мы получаем зависимости между координатами точек $M(x, y)$ и $M'(x', y')$:

$$\begin{cases} x'=\lambda x \\ y'=\lambda y \end{cases} \quad \text{и} \quad \begin{cases} x=\dfrac{x'}{\lambda}, \\ y=\dfrac{y'}{\lambda}. \end{cases} \tag{52}$$

Совершенно очевидно, что если обозначим через $\{M_n(x_n, y_n)\}$ последовательность точек фигуры F, имеющую пределом точку $M(x, y)$ той же фигуры, и через $\{M'(x'_n, y'_n)\}$ и $M'(x, y)$ — гомологические точки фигуры F', то взаимная непрерывность отображения выразится в том, что если

$$x_n \to x \quad \text{и} \quad y_n \to y,$$

то

$$x'_n \to x' \quad \text{и} \quad y'_n \to y',$$

и обратно.

Другой, менее тривиальный, пример гомеоморфного отображения — инверсия фигуры F на фигуру F' (см. § 56, стр. 262); формулы (52) заменяются теперь следующими:

$$\begin{cases} x' = \dfrac{x}{x^2+y^2}, \\ y' = \dfrac{y}{x^2+y^2}, \end{cases} \quad \text{и} \quad \begin{cases} x = \dfrac{x'}{x'^2+y'^2}, \\ y = \dfrac{y'}{x'^2+y'^2}. \end{cases} \tag{53}$$

Если фигура (множество точек) F не содержит начала координат, то взаимная непрерывность, как легко видеть из формул (53), обеспечена.

Прежде чем перейти к дальнейшим примерам гомеоморфных отображений (причём мы ограничимся примерами на плоскости с обычной метрикой), установим одно вспомогательное предложение общего характера.

Если множество $\mathcal{X}$ гомеоморфно отображено на множество $\mathcal{Y}$, то тем самым всякая часть $\mathcal{X}_1$ множества $\mathcal{X}$ также гомеоморфно отображается на ту часть $\mathcal{Y}_1$ множества $\mathcal{Y}$, которая ей соответствует в силу установленного отображения.

В самом деле, если $\{X_n\}$ — последовательность элементов из $\mathcal{X}_1$, сходящаяся к элементу X, тоже из $\mathcal{X}_1$, то

$$f(X_n) \to f(X),$$

где все элементы $Y_n = f(X_n)$ $(n = 1, 2, 3, \ldots)$ и элемент $Y = f(X)$ принадлежат множеству $\mathcal{Y}$. Но так как каждый из них по построению множества $\mathcal{Y}_1$ принадлежит также и этому последнему множеству и так как всё это рассуждение в целом обратимо, то соотношение, установленное между $\mathcal{X}_1$ и $\mathcal{Y}_1$, является гомеоморфным отображением.

Таким образом, если

$$\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2 \quad \text{и} \quad \mathcal{X}_1\mathcal{X}_2 = 0$$

и если $\mathcal{X}$ гомеоморфно отображено на $\mathcal{Y}$, причём $\mathcal{X}_1$ отображено на $\mathcal{Y}_1$, а $\mathcal{X}_2$ — на $\mathcal{Y}_2$, то оба отображения $\mathcal{X}_1$ на $\mathcal{Y}_1$ и $\mathcal{X}_2$ на $\mathcal{Y}_2$ — тоже гомеоморфные.

В частности, допуская, что $\mathcal{X}_1$ состоит из одного элемента, приходим к заключению: *если множество $\mathcal{X}$ гомеоморфно отображено на множество $\mathcal{Y}$, то отображение между множествами, полученными после удаления из $\mathcal{X}$ одного элемента и соответственного ему элемента из $\mathcal{Y}$, также продолжает оставаться гомеоморфным.*

Этот результат совместно с теоремами предыдущего параграфа позволит нам в ряде случаев убедиться в невозможности установить между двумя множествами гомеоморфное отображение.

Всякие два замкнутых отрезка гомеоморфны между собой (так как они подобны). Также и всякие два открытых отрезка (без конечных точек). Но замкнутый и открытый (или полузамкнутый) отрезки

не гомеоморфны, так как при непрерывном отображении замкнутого множества получается также замкнутое множество (§ 61, теорема II). Полузамкнутый отрезок и открытый отрезок также не гомеоморфны: действительно, допуская противное, удалим из полузамкнутого отрезка его единственную концевую точку и из открытого — ту точку, которая гомологична этой точке; оставшиеся множества по предыдущему должны быть гомеоморфны; но это невозможно, так как одно из них связно, другое же — наверное не связно (§ 61, теорема I).

Один отрезок не гомеоморфен паре (объединению) отрезков без общих точек (по той же теореме). Один отрезок не гомеоморфен также системе (объединению) трёх отрезков без общих точек; пара отрезков без общих точек не гомеоморфна системе трёх отрезков без общих точек и т. п. Но один замкнутый отрезок гомеоморфен паре замкнутых отрезков, имеющих общую концевую точку. Например, отрезок $[0, 1]$ гомеоморфен ломаной линии ABC. Для доказательства достаточно установить хотя бы одно гомеоморфное отображение; можно отобразить отрезок $\left[0, \frac{1}{2}\right]$ на AB посредством подобия и отрезок $\left[\frac{1}{2}, 1\right]$ на BC таким же образом. Вообще любая ломаная линия, себя не пересекающая и не образующая замкнутого многоугольника, гомеоморфна отрезку. Таковы, например, буквы [1])

Г, М, П.

Однако буква Т не гомеоморфна отрезку. В самом деле, допуская противное, удалим из этой буквы среднюю точку горизонтального отрезка, являющуюся вместе с тем концевой для вертикального, а из отрезка, гомеоморфного рассматриваемой букве, — точку, гомологичную удалённой: останется в одном случае система из трёх отрезков без общих точек, в другом — один отрезок или пара отрезков.

Буквы

Е, Ш, Ч, Ц

также не гомеоморфны отрезку, но гомеоморфны букве Т и, следовательно, гомеоморфны между собой.

Буква Щ гомеоморфна букве Н, но не гомеоморфна ни отрезку, ни какой-либо из ранее названных букв.

Буква О (окружность) не гомеоморфна ни отрезку, ни ранее названным буквам. Действительно, при противоположном допущении, удаляя из отрезка какую-нибудь точку, отличную от концевых (если

[1]) Разумеется, «буквы» понимаются здесь в идеализированном смысле — как бы проведённые «бесконечно тонким» пером.

таковые имеются) и гомологичную ей точку окружности, мы получили бы в одном случае несвязное, в другом — связное множество.

Эллипс гомеоморфен окружности: растяжение, как и его частный случай — преобразование подобия, представляет собой гомеоморфное отображение.

Читатель легко продлит перечень примеров этого рода. Может быть, занимаясь «буквами», он исчерпает весь русский алфавит, разбив буквы на ряд групп таких, что буквы из одной группы будут гомеоморфны между собой, а из различных групп — не гомеоморфны.

Дальнейшие и более разнообразные примеры содержатся в статьях, посвящённых топологии (см. Э. э. м., кн. IV).

Здесь же мы имеем в виду остановиться ещё на одной стороне вопроса. Ограничиваясь простейшим примером гомеоморфного отображения замкнутого отрезка на замкнутый отрезок, постараемся выяснить, в какой степени неопределённой является задача реализовать (установить) такого рода отображение.

Мы уже видели раньше (§ 52), что всякая монотонная (возрастающая или убывающая) непрерывная функция $y=f(x)$ устанавливает гомеоморфное отображение отрезка $[a, b]$ *на оси* Ox на отрезок $[A, B]$ на оси Oy.

Убедимся в справедливости обратного утверждения: *если между отрезками* $[a, b]$ *на оси* Ox *и* $[A, B]$ *на оси* Oy *установлено гомеоморфное отображение, то оно реализуется посредством функции* $y=f(x)$, *обладающей свойствами монотонности и непрерывности.*

Поскольку гомеоморфное отображение есть частный случай о д н о з н а ч н о г о, можно заключить о с у щ е с т в о в а н и и функции $y=f(x)$, реализующей данное отображенне; поскольку данное гомеоморфное отображение устанавливает н е п р е р ы в н у ю зависимость между точками отрезка $[a, b]$ и точками отрезка $[A, B]$, можно судить о н е п р е р ы в н о с т и функции $f(x)$. Остаётся убедиться в её м о н о т о н н о с т н.

Концевой точке одного отрезка должна непременно соответствовать концевая же точка другого отрезка: иначе, если бы, например, точке a первого отрезка соответствовала внутренняя точка η второго отрезка, то, удаляя эти обе точки, мы получили бы в одном случае связное, в другом — несвязное множество. Возможны два случая: 1) $f(a)=A$, $f(b)=B$, 2) $f(a)=B$, $f(b)=A$; остановимся на первом.

Так как предполагается, что $a<b$, $A<B$, то функция $f(x)$ не может быть убывающей; докажем, что она — возрастающая. При противоположном допущении существовали бы такие две точки x' и x'', что

$$a \leqslant x' < x'' \leqslant b$$

и вместе с тем было бы

$$f(x') > f(x'').$$

Допустим сначала, что $a<x'$. Возможность равенства $f(a)=f(x')$ исключена вследствие взаимной однозначности отображения.

Предположим, что $f(a)<f(x')$. Пусть C — какое-нибудь число, заключённое между наибольшим из чисел $f(a)$ и $f(x'')$ и числом $f(x')$. По свойству I непрерывных функций (§ 47) функция $f(x)$ должна принимать хоть раз значение C в промежутке (a, x') и в промежутке (x', x''); а это несовместно со взаимной однозначностью отображения.

Если бы было $f(a)>f(x')$, то аналогичным образом пришлось бы рассуждать относительно промежутков (x', x'') и (x'', b)[1]), взяв в качестве C

[1]) Равенство $x''=b$ не было бы возможно, так как мы имели бы $f(x'')<f(x')<f(a)$ и вместе с тем $f(a)<f(b)$.

число, заключённое между $f(x'')$ и наименьшим из чисел $f(x')$ и $f(b)$. Это же самое рассуждение было бы пригодно и при допущении $a = x'$.

Так как случай 2) вполне аналогичен случаю 1), то теорему следует считать доказанной.

§ 63. Верхняя и нижняя границы числовых множеств или последовательностей. Верхний и нижний пределы числовых множеств или последовательностей

Среди прочих множеств числовые множества (т. е. множества действительных чисел) резко выделяются свойством быть «упорядоченными» естественным образом — по величине: относительно каждых двух различных элементов числового множества x и y всегда можно сказать, которое из двух соотношений порядка

$$x < y \text{ или } x > y$$

будет справедливым.

Мы пользуемся в дальнейшем обыкновенной метрикой, в которой

$$\rho(x, y) = |x - y|.$$

Пусть множество E ограничено сверху: это значит, что существует такое число M, что всякое число x из E меньше, чем M:

$$x < M. \tag{54}$$

Если множество E ограничено сверху и число M_0 обладает свойствами:

1° $x \leqslant M_0$ для всякого элемента x из E,

2° $x = M_0$ хотя бы для одного элемента x_0 из E,

то число M_0 есть наибольший элемент из E (иначе, максимум).

Не всякое ограниченное множество содержит наибольший элемент. Об этом свидетельствуют примеры, отчасти нам уже знакомые: открытый справа промежуток (0, 1) или [0, 1), множество правильных рациональных дробей, множество чисел вида $1 - \frac{1}{n}$ или множество чисел вида $1 - \frac{1}{m} - \frac{1}{n}$ (где m и n — натуральные числа) и т. п.

Число G, связанное с данным ограниченным сверху множеством E, называется *верхней границей* (или *верхней гранью*) множества E, если оно обладает двумя свойствами:

1° $x \leqslant G$ для всякого элемента x из E,

2° как бы мало ни было $\varepsilon\ (> 0)$, $x > G - \varepsilon$ хотя бы для одного элемента x из E.

Очевидно, множество E не может иметь двух различных верхних границ. В самом деле, если бы E имело верхние границы G_1 и G_2, причём было бы, например, $G_1 < G_2$, то для некоторого элемента мы должны были бы иметь неравенства

$$x > G_2 - \varepsilon \quad \text{и} \quad x \leqslant G_1,$$

которые, однако, при выборе ε согласно условию $\varepsilon < G_2 - G_1$ содержали бы противоречие.

Если множество E имеет наибольший элемент M_0, то этот наибольший элемент M_0 и есть верхняя граница E:

$$G = M_0.$$

Действительно, условие 1° одно и то же для верхней границы и для наибольшего элемента; что же касается условия 2°, то как бы мало ни было $\varepsilon\ (>0)$, из равенства $x = M_0$, конечно, следует неравенство $x > M_0 - \varepsilon$.

Теорема. *Всякое непустое ограниченное сверху числовое множество E имеет верхнюю границу.*

Пусть M — такое число, что для всех x из E справедливо неравенство $x < M$, и L — число, меньшее чем какой-нибудь элемент множества E.

Тогда промежуток $[L, M]$ длины $\Delta = M - L$ обладает следующими свойствами:

(α) в нём содержится хоть один элемент множества E;

(β) нет ни одного элемента множества E, который был бы больше любого числа из промежутка (который, другими словами, был бы больше, чем его правый конец).

Рассмотрим теперь промежутки $\left[L, \frac{L+M}{2}\right]$ и $\left[\frac{L+M}{2}, M\right]$. Один из них (по меньшей мере) обладает свойствами (α) и (β); обозначим этот промежуток через $[L_1, M_1]$.

Рассмотрим далее промежутки $\left[L_1, \frac{L_1+M_1}{2}\right]$ и $\left[\frac{L_1+M_1}{2}, M_1\right]$. Один из них (по меньшей мере) обладает свойствами (α) и (β); обозначим его через $[L_2, M_2]$.

Продолжая таким образом до бесконечности, мы получим последовательность промежутков $\{[L_n, M_n]\}$, обладающих каждый свойствами (α) и (β), и таких, что

1) каждый промежуток $[L_n, M_n]$ содержит следующий за ним $[L_{n+1}, M_{n+1}]$,

2) длины промежутков образуют сходящуюся геометрическую прогрессию $M_{n+1} - L_{n+1} = \frac{1}{2}(M_n - L_n)$, так что $M_n - L_n = \frac{\Delta}{2^n}$.

Отсюда следует существование общего предела, который мы обозначим через G: $\lim L_n = \lim M_n = G$.

Очевидно, при любом n справедливы неравенства $L_n \leqslant G \leqslant M_n$.

Покажем, что число G обладает свойствами 1° и 2°.

Пусть существует такой элемент x_0 из E, что $x_0 > G$. Тогда, выбрав n по условию $\frac{\Delta}{2^n} < x_0 - G$, мы убедимся, что $x_0 > G + \frac{\Delta}{2^n} \geqslant L_n + \frac{\Delta}{2^n} = M_n$, а это противоречит свойству (β) промежутка $[L_n, M_n]$. Итак, число G обладает свойством 1°.

Пусть теперь ε — некоторое положительное число и пусть все элементы x из E удовлетворяют неравенству $x \leqslant G - \varepsilon$. В таком случае, выбрав n по условию $\frac{\Delta}{2^n} < \varepsilon$, мы получим опять для всех элементов множества E:

$$x < G - \frac{\Delta}{2^n} \leqslant M_n - \frac{\Delta}{2^n} = L_n,$$

а это противоречит свойствам (α) промежутка $[L_n, M_n]$. Итак, число G обладает свойством 2°.

Но раз число G обладает свойствами 1° и 2° верхней границы, значит, оно и есть верхняя граница нашего множества E.

Определение верхней границы G множества E, ограниченного сверху, можно перефразировать следующим образом.

Верхняя граница G множества E есть одно из двух:

если существует наибольший элемент множества, то G с ним совпадает;

если не существует наибольшего элемента множества E, то G есть наименьшее из чисел M, удовлетворяющих неравенству (54), где x — произвольный элемент из E.

В первом случае само число G, очевидно, есть элемент множества E; во множестве же чисел M, удовлетворяющих неравенству (54), не существует наименьшего [1]).

Во втором случае G не есть элемент множества E, иначе неравенство $x < M\,(= G)$ должно было бы удовлетворяться при $x = G$.

Можно утверждать также и следующее:

Верхняя граница G множества E, ограниченного сверху, есть одно из двух:

или изолированный наибольший элемент множества E,

или предельная точка (см. § 36) множества E.

(*Изолированным элементом* множества мы называем такой элемент, что в некоторой его окрестности нет других элементов множества.)

В самом деле, пусть G есть наибольший элемент E, но не изолированный. Возьмём последовательность положительных чисел $\{\varepsilon_n\}$, удовлетворяющую требованию $\varepsilon_n \to 0$, и, пользуясь свойством 2° верхней границы, к каждому ε_n подберём элемент x_n множества E по условиям $x_n > G - \varepsilon_n$ и $x_n \neq G$ [2]). Вместе с тем по свойству 1° $x_n \leqslant G$, и потому, как легко понять, $x_n \to G$. Среди точек последовательности $\{x_n\}$ имеется бесконечное множество различных; действительно, имея уже точку x_{n_1}, мы получим точку x_{n_2}, наверное отличную от x_{n_1}, если n_2 будет настолько велико, что $\varepsilon_{n_2} < G - x_{n_1}$,

---

[1]) Любое число, большее чем G, может быть взято в качестве M; среди этих чисел, однако, нет наименьшего.

[2]) Такой выбор возможен, так как G — не изолированный элемент E.

и т. д. Отсюда ясно, что в любой окрестности точки G имеется бесконечное множество различных точек E, т. е. G есть предельная точка множества.

Отнюдь не исключено, конечно, что верхняя граница G множества есть одновременно и наибольший его элемент (не изолированный) и его предельная точка. Примерами могут служить: точка 1 в случае промежутка $[0, 1]$ или та же точка в случае множества, составленного из точек вида $1 - \frac{1}{n}$ (n — натуральное) и ещё точки 1.

Всякое замкнутое множество E содержит наибольший элемент. Именно наибольшим элементом замкнутого множества E является его верхняя граница G.

В самом деле, мы только что доказали, что G есть или изолированный наибольший элемент множества E или его предельная точка. Но если G есть предельная точка E, то G есть элемент производного множества E' и раз множество E' замкнуто (т. е. $E' \subseteq E$, см. § 60), то G должно быть также элементом E.

Мы можем теперь обратиться к понятию верхнего предела множества, ограниченного сверху.

Число L, связанное с данным множеством E, ограниченным сверху, называется *верхним пределом* этого множества, если оно обладает свойствами:

1° Как бы мало ни было $\varepsilon\,(>0)$, неравенству $x < L + \varepsilon$ удовлетворяют «почти все» элементы E (т. е. все, кроме, может быть, конечного их числа).

2° Как бы мало ни было $\varepsilon\,(>0)$, неравенству $x > L - \varepsilon$ удовлетворяет бесконечное множество элементов E.

Конечные множества, очевидно, верхнего предела не имеют.

В случае же, если ограниченное сверху множество E бесконечно, свойства 1° и 2° верхнего предела L множества E можно охватить следующей строго эквивалентной формулировкой:

Верхний предел L множества E есть наибольшая из предельных точек этого множества (т. е. наибольший элемент производного множества E').

Действительно, из условий 1° и 2° вытекает немедленно, что в любой окрестности точки L содержится бесконечное множество элементов множества E, так что L есть предельная точка E. Вместе с тем L — наибольшая из предельных точек E, так как существование ещё одной предельной точки L', большей чем L, противоречило бы условию 1° (если бы ε было выбрано меньшим, чем $L' - L$).

Обратно, если L есть наибольшая из предельных точек E, то условие $x < L + \varepsilon$ может не быть выполнено лишь для конечного множества точек из E: иначе по теореме Больцано-Вейерштрасса (см. § 36, примечание 2) множество E имело бы хоть одну предельную точку L' такую, что $L' \geqslant L + \varepsilon$. Таким образом мы приходим к свой-

ству 1°. Свойство же 2° следует из того, что L есть предельная точка E.

Теорема. *Всякое ограниченное сверху бесконечное числовое множество E имеет верхний предел.*

В самом деле, производное множество E' непременно замкнуто (см. § 37, стр. 163); следовательно по предыдущему оно имеет наибольший элемент.

Верхний предел бесконечного множества не превышает его верхней границы:

$$L \leqslant G. \tag{55}$$

В противном случае, т. е. при допущении $L > G$, свойство 2° верхнего предела и свойство 1° верхней границы оказались бы в противоречии.

Легко понять, что в соотношении (55) должен стоять знак неравенства или знак равенства, смотря по тому, является ли верхняя граница изолированным элементом множества E или его предельной точкой.

Следующие понятия, отношения и теоремы вполне аналогичны изложенным выше.

Условие

$$x > m, \tag{56}$$

предполагаемое выполненным для всех элементов множества E, означает, что это множество ограничено снизу.

Если число m_0 таково, что

1° $x \geqslant m_0$ для всякого элемента x из E,

2° $x = m_0$ хотя бы для одного элемента x_0 из E,

то оно называется *наименьшим элементом* E (иначе, *минимумом*).

Число g называется *нижней границей* (или *нижней гранью*) множества E, если справедливо следующее:

1° $x \geqslant g$ для всякого элемента x из E,

2° как бы мало ни было ε (>0), $x < g + \varepsilon$ хотя бы для одного элемента x_0 из E.

Если множество E имеет наименьший элемент m_0, то m_0 есть вместе с тем нижняя граница g:

$$g = m_0. \tag{57}$$

Теорема. *Всякое непустое ограниченное снизу числовое множество имеет нижнюю границу.*

Нижняя граница есть одно из двух:

или наименьший элемент множества E,

или наибольшее из чисел m, удовлетворяющих неравенству (56) для всех x из E.

Вместе с тем g есть одно из двух:

или изолированный наименьший элемент множества E,

или предельная точка этого множества.

Всякое замкнутое множество содержит наименьший элемент.

Число l называется *нижним пределом* множества E, если справедливо следующее:

1° При сколь угодно малом ε (>0) неравенству $x>l-\varepsilon$ удовлетворяют «почти все» элементы E.

2° При сколь угодно малом ε (>0) неравенству $x<l+\varepsilon$ удовлетворяет бесконечное множество элементов E.

Иначе говоря (для случая бесконечных множеств):
l есть наименьшая из предельных точек множества E.

Теорема. *Всякое ограниченное снизу бесконечное множество имеет нижний предел.*

Нижний предел бесконечного множества больше или равен его нижней границе:

$$l \geqslant g. \tag{58}$$

Так как вполне очевидно для бесконечных множеств, что никогда нижний предел не превышает верхнего, то мы получаем цепь неравенств:

$$g \leqslant l \leqslant L \leqslant G. \tag{59}$$

Случай $l=L$ возможен тогда и только тогда, когда множество имеет только одну предельную точку. Такая точка является просто *пределом* множества.

Случай $g=G$ возможен тогда и только тогда, когда множество состоит из одной точки.

Мы рассматривали до сих пор только множества, ограниченные сверху или снизу. Если множество не ограничено сверху, иногда условливаются полагать, что

$$L=G=+\infty.$$

Если множество не ограничено снизу, полагают таким же образом:

$$l=g=-\infty.$$

Понятия верхней и нижней границ множества и верхнего и нижнего пределов множества переносятся также на случай последовательности.

Имея некоторую последовательность чисел $\{a_n\}$, мы всегда можем связать с нею множество E чисел, являющихся членами этой последовательности.

Под верхней и нижней границей последовательности понимают (соответственно) не что иное, как верхнюю и нижнюю границу связанного с нею множества — невзирая на то, что это множество может оказаться и конечным. Таким образом, определение верхней, например, границы последовательности почти буквально совпадает с определением верхней границы множества:

Число G называется *верхней границей (гранью) последовательности* $\{a_n\}$, если оно обладает свойствами:

1° $a_n \leqslant G$ при всех значениях n ($n = 1, 2, 3, \ldots$),

2° как бы мало ни было $\varepsilon\,(>0)$, $a_n > G - \varepsilon$ хотя бы при одном значении n.

Легко понять, что верхняя граница последовательности совпадает с верхней границей множества элементов этой последовательности.

Что касается определения верхнего (и соответственно нижнего) предела последовательности, то, хотя оно формулируется почти буквально так же, как для множества, в содержании этого понятия есть некоторая особенность.

Число L называется *верхним пределом последовательности* $\{a_n\}$, ограниченной сверху, если оно обладает свойствами:

1° как бы мало ни было $\varepsilon\ (>0)$, неравенство $a_n < L + \varepsilon$ выполняется при «почти всех» значениях n (т. е. для достаточно больших значений n),

2° как бы мало ни было $\varepsilon\,(>0)$, неравенство $a_n > L - \varepsilon$ выполняется для бесконечного множества различных значений n (т. е. для сколь угодно больших значений n).

При таких условиях оказывается, что *любая последовательность, ограниченная сверху, имеет верхний предел* даже в том случае, если множество, составленное из членов последовательности, содержит лишь конечное число элементов.

Аналогично, конечно, для нижнего предела.

Так, верхний и нижний пределы последовательности $\{(-1)^n\}$ соответственно равны $+1$ и -1, хотя множество членов последовательности конечно: состоит всего из двух элементов — этих же самых чисел $+1$ и -1[1]).

Считаем нужным познакомить читателя с обозначениями для верхней и нижней границы множества E или последовательности $\{a_n\}$:

$$G = \sup E, \quad g = \inf E \text{ [2])}$$

и

$$G = \sup \{a_n\}, \quad g = \inf \{a_n\}.$$

Что касается верхнего и нижнего предела, то для них приняты обозначения:

$$L = \overline{\lim} E, \quad l = \underline{\lim} E$$

и

$$L = \overline{\lim}\, a_n, \quad l = \underline{\lim}\, a_n.$$

---

[1]) Сопоставляя понятия множества чисел и последовательности чисел, не следует упускать из виду и то обстоятельство (в данной связи нас не интересующее), что *не всякое множество чисел может быть расположено в форме последовательности*; это невозможно для случая множеств несчётных (см. Э. э. м., кн. I, стр. 93).

[2]) «sup» и «inf» — сокращения латинских слов «superior» и «inferior» — «верхний» и «нижний».

Так, в случае множества или[1]) последовательности чисел $\left\{(-1)^n \frac{n+1}{n}\right\}$ (n — натуральное) мы получаем:

$$G = \sup\left\{(-1)^n \frac{n+1}{n}\right\} = \frac{3}{2}, \quad g = \inf\left\{(-1)^n \frac{n+1}{n}\right\} = -2,$$

$$L = \overline{\lim}\,(-1)^n \frac{n+1}{n} = 1, \qquad l = \underline{\lim}\,(-1)^n \frac{n+1}{n} = -1.$$

Запись $\overline{\lim}\, a_n = L$, как легко понять, может быть истолкована в том смысле, что для некоторой последовательности натуральных чисел $\{n_p\}$ имеет место предельное отношение (в обычном смысле) $\lim a_{n_p} = L$, тогда как при любом $\varepsilon\,(>0)$ выбрать последовательность $\{n_p\}$, удовлетворяющую требованию $\lim a_{n_p} = L + \varepsilon$, уже невозможно.

В соответствии с этим запись

$$\overline{\lim_{x \to c}}\, f(x) = L \tag{60}$$

понимается в том смысле, что существует последовательность $\{x_n\}$, $x_n \to c$, для которой $\lim f(x_n) = L$, тогда как при сколь угодно малых ε выбрать последовательность $\{x_n\}$, $x_n \to c$, удовлетворяющую требованию $\lim f(x_n) = L + \varepsilon$, уже невозможно.

Аналогично — относительно нижнего предела.

Числа $\overline{\lim\limits_{x \to c}}\, f(x)$ и $\underline{\lim\limits_{x \to c}}\, f(x)$ носят названия *верхнего и нижнего предела значений функции* $f(x)$ *в точке* $x = c$.

Очевидно, что

$$\underline{\lim_{x \to c}}\, f(x) \leqslant \overline{\lim_{x \to c}}\, f(x),$$

причём знак равенства достигается в том и только в том случае, если функция $f(x)$ непрерывна в точке $x = c$.

Примером может служить функция $f(x) = \sin\frac{1}{x}$ (см. § 41, стр. 182); для неё мы получаем:

$$\overline{\lim_{x \to 0}}\, \sin\frac{1}{x} = +1, \quad \underline{\lim_{x \to 0}}\, \sin\frac{1}{x} = -1.$$

Вводя дополнительные ограничения $x > c$ или $x < c$, можно также определить понятия правого верхнего и правого

[1]) В данном случае выбор термина безразличен, так как среди членов последовательности нет одинаковых.

нижнего, а также левого верхнего и левого нижнего предела значений функции в данной точке.

В заключение следует обратить внимание читателя на то, что все введённые выше понятия, в первую очередь — понятия верхней и нижней границы множества, нередко служат целям сокращения и упрощения («унификации») доказательств.

Приведём несколько примеров, причём ограничимся теоремами, нам уже известными. Намечаем лишь схему доказательств, предлагая читателю в качестве упражнения воспроизвести их детали.

1. *Любое ограниченное сверху бесконечное множество имеет верхний предел* (см. стр. 291).

Пусть E — данное множество. Рассмотрим множество $\mathscr{E}$ точек x, из которых каждая меньше лишь конечного числа точек множества E. Нижняя граница множества $\mathscr{E}$ и есть верхний предел множества E.

2. *Если функция $f(x)$ непрерывна в промежутке $[a, b]$, причём $f(a) < 0$, $f(b) > 0$, то в некоторой внутренней точке промежутка $f(x)$ обращается в нуль* (теорема Больцано, § 47, стр. 215).

Рассмотрим множество $\mathscr{E}$ точек x, в которых $f(x) > 0$. Нижняя граница множества $\mathscr{E}$ представляет собой точку, в которой функция $f(x)$ обращается в нуль.

3. *Если функция $f(x)$ непрерывна в промежутке $[a, b]$, то множество её значений в этом промежутке имеет наибольший элемент* (теорема Вейерштрасса, § 47, стр. 218).

Пусть $\mathscr{E}$ — множество значений функции $f(x)$ в рассматриваемом промежутке, G — его верхняя граница. Допустим (доказывая от противного), что не существует такой точки $x = \xi$, в которой $f(\xi) = G$.

Рассмотрим новую функцию

$$\varphi(x) = \frac{1}{G - f(x)}.$$

Она непрерывна в промежутке $[a, b]$ и потому (теорема II на стр. 217) ограничена сверху:

$$\varphi(x) < M.$$

В таком случае при всех значениях x из промежутка

$$f(x) < G - \frac{1}{M},$$

и следовательно, G не есть верхняя граница множества значений $f(x)$.

4. *Замкнутый промежуток $\mathscr{A} \equiv [a, b]$ — связное множество* (§ 60, стр. 278). Пусть $\mathscr{A}$ не связно; тогда существуют такие множества $\mathscr{M}$ и $\mathscr{N}$, что

$$1^\circ\ \mathscr{A} = \mathscr{M} + \mathscr{N},$$

$$2^\circ\ \mathscr{M}\mathscr{N}' + \mathscr{M}'\mathscr{N} + \mathscr{M}'\mathscr{N}' = 0,$$

$$3^\circ\ \mathscr{M} \neq 0,\ \mathscr{N} \neq 0.$$

Предположим, например, что точка a принадлежит множеству $\mathscr{M}$; обозначим через g нижнюю границу множества $\mathscr{N}$. Точка g или принадлежит $\mathscr{N}$ или не принадлежит; но в последнем случае она принадлежит множеству $\mathscr{M}$.

Если g принадлежит $\mathscr{N}$, то все точки x, удовлетворяющие неравенству $a \leqslant x < g$, принадлежат $\mathscr{M}$; тогда само g есть предельный элемент $\mathscr{M}$, так что $\mathscr{M}'\mathscr{N} \neq 0$, что противоречит условию 2°.

Если g принадлежат $\mathscr{M}$, то во множестве $\mathscr{N}$ существуют точки, сколь угодно близкие к g, и значит, g есть предельный элемент $\mathscr{N}$, так что $\mathscr{M}\mathscr{N}' \neq 0$, а это противоречит условию 2°.

Теорема доказана [1]).

Предоставляем читателю судить о том, насколько выигрывают доказательства предшествующих теорем в отношении чёткости и сжатости благодаря использованию понятия границ.

---

[1]) См. Э. э. м., кн. IV, П. С. Александров и В. А. Ефремович, Основные топологические понятия, гл. 3, п. 3.

И. П. НАТАНСОН

ПРОИЗВОДНЫЕ, ИНТЕГРАЛЫ И РЯДЫ

ВВЕДЕНИЕ

Величины, с которыми людям постоянно приходится иметь дело при изучении природы, являются в большинстве случаев величинами изменяющимися или, как часто говорят, переменными. Температура воздуха, давление пара в котле, напряжение тока в электрической сети, скорость самолёта, — все эти величины с течением времени изменяются и, следовательно, являются величинами переменными. Однако лишь в эпоху XVI — XVII вв. под влиянием практических потребностей бурно развивающегося естествознания и техники математика овладела общим понятием переменной величины, и с этих пор переменные величины стали основным объектом математических исследований. Большую роль при этом сыграли математические работы Декарта.

«Поворотным пунктом, — говорит Ф. Энгельс, — в математике была декартова *переменная величина.* Благодаря этому в математику вошли *движение и диалектика* и благодаря этому же стало *немедленно необходимым дифференциальное и интегральное исчисление*, которое тотчас и возникает и которое было в общем и целом завершено, а не изобретено, Ньютоном и Лейбницем» [1]).

Можно сказать кратко, что математический анализ — это математика величин переменных.

Для более полной характеристики предмета математического анализа следует указать, что он занимается изучением переменных величин, рассматривая их не изолированно, а в их взаимной связи. Точным математическим понятием, выражающим идею взаимосвязи переменных величин, является понятие функции. Это есть основное и важнейшее понятие математического анализа.

Идеи математического анализа, а именно, идеи переменной величины и функции имеют чрезвычайно важное значение для элементарной математики. Вся теория тригонометрических функций представляет, по сути дела, элементарную главу математического анализа. В школьном курсе алгебры изучаются рациональные, простейшие алгебраические, иррациональные функции, а также и некоторые неалгебраические (трансцендентные) функции: таковы степенная

---

[1]) Ф. Энгельс, Диалектика природы, Госполитиздат, 1948, стр. 208.

функция с иррациональным показателем, показательная, логарифмическая; здесь же изучаются основы теории пределов и простейший ряд — бесконечная геометрическая прогрессия. В геометрии размеры криволинейных фигур и тел, ограниченных кривыми поверхностями, определяются как пределы размеров некоторых других фигур и тел, изменяющихся и приближающихся к данным.

К числу важнейших разделов математического анализа, имеющих разнообразные приложения во всех областях математики, физики и техники, относятся дифференциальное и интегральное исчисление, а так же теория рядов. Разделы эти не представлены в настоящее время в программе общеобразовательной школы. Однако понятия производной и интеграла давно уже относятся к основным понятиям науки, имеют огромную образовательную ценность и со временем могут оказаться включёнными (разумеется, в очень скромном объёме) в курс математики средней школы.

Настоящая статья содержит в систематическом виде основные сведения о производных, интегралах и рядах. Сведения эти представлены, разумеется, в объёме, во много раз превышающем тот, в котором они могут войти в школьный курс. Выбор этих сведений здесь, как и в других разделах «Энциклопедии элементарной математики», определяется потребностями изучения и исследования элементарных функций, их вычисления и решения ряда задач геометрии и физики (проведение касательной к кривой, определение скорости по пути, или пути по скорости, определение площади плоских фигур, длин кривых, объёмов и площадей поверхностей простейших тел).

Не ставя своей целью дать хоть сколько-нибудь полный очерк истории развития математического анализа, отметим лишь некоторые моменты этой истории. Наличие некоторых важных идей анализа можно усмотреть уже в античной науке. Классический «метод исчерпывания» представляет собой прообраз теории рядов. В некоторых работах Архимеда применяется в зачаточной форме ход мыслей, характерный для позднейшего интегрального исчисления.

Примерно с середины XVII в. достижения древнегреческой науки в области математики были уже превзойдены и, как уже упоминалось, в конце XVII в. трудами Ньютона (1642—1727) и Лейбница (1646—1716) было завершено построение дифференциального и интегрального исчисления.

Говоря о «завершении» процесса создания математического анализа, мы имеем в виду установление основных руководящих принципов анализа и, в первую очередь, установление взаимной обратности задач дифференциального исчисления и интегрального исчисления. Было бы неправильно понять термин «завершение» в том смысле, что после работ Ньютона и Лейбница развитие анализа остановилось. Наоборот, научное творчество в этой области интен-

сивно продолжалось и в XVIII и в XIX вв. и весьма успешно протекает и поныне. В настоящее время по различным отделам математического анализа публикуется каждый год более 1000 работ.

Начало разработки проблем анализа у нас в России связано с именем знаменитого Л. Эйлера (1707—1783). Эйлер, швейцарец по происхождению, почти всю свою жизнь провёл в Петербурге, был членом Петербургской Академии наук. Он занимался самыми разнообразными вопросами математики и механики. В частности, ему принадлежат важные заслуги в области дифференциального и интегрального исчисления.

Ряд важных открытий в интегральном исчислении (техника интегрирования, теория кратных интегралов, дифференциальные уравнения, вариационное исчисление) был сделан академиком Михаилом Васильевичем Остроградским (1801—1861). Его знаменитая формула для преобразования кратных интегралов и принадлежащий ему способ интегрирования рациональных дробей вошли во все учебные руководства.

Кроме Остроградского, вопросами анализа занимался и другой современный ему петербургский академик — Виктор Яковлевич Буняковский (1804—1889), опубликовавший в этой области ряд работ. Интересные исследования по математическому анализу выполнил и наш великий геометр Николай Иванович Лобачевский (1792—1856).

Создателем большой научной школы математического анализа был гениальный русский математик Пафнутий Львович Чебышев (1821—1894). Он — автор ряда важных работ по интегральному исчислению (как по вопросам интегрирования в элементарных функциях, так и по приближённому вычислению определённых интегралов), но главной заслугой Чебышева в области анализа явились его выдающиеся исследования по теории приближения функций. Дальнейшее развитие этих исследований привело к созданию новой важной ветви анализа — конструктивной теории функций, имеющей большое прикладное значение.

Будучи в течение ряда лет профессором Петербургского университета, Чебышев создал мощную школу, наиболее яркими представителями которой были академики А. А. Марков (1856—1922) и А. М. Ляпунов (1857—1918). Традиции этой школы и по сие время живы в Ленинградском университете.

В XX в. главным образом под влиянием Н. Н. Лузина (1883—1950) и Д. Ф. Егорова (1869—1931) складывается московская математическая школа. Полного расцвета её деятельность достигает уже в советский период и в настоящее время по широте своих интересов и важности результатов московская школа бесспорно занимает первое место в мире. Выдающимися её представителями в области анализа являются А. Н. Колмогоров, И. Г. Петровский, И. И. Привалов, Д. Е. Меньшов и др.

Одним из крупнейших современных математиков является академик С. Н. Бернштейн (р. 1880), продолжатель исследований П. Л. Чебышева по теории приближения функций (а также и по теории вероятностей). Под влиянием С. Н. Бернштейна воспитался целый ряд молодых советских исследователей.

Колоссальный общий подъём культуры и науки, характерный для страны победившего социализма, ярко проявляется и в области математики. Выше уже было отмечено выдающееся значение деятельности московской математической школы. Весьма значительные и обширные математические исследования ведутся и в других культурных центрах нашей страны — Ленинграде, Киеве, Харькове, Одессе, Тбилиси, Казани, Ереване и др.

Обстоятельное изложение успехов математических наук в СССР (до 1947 г.) можно найти в коллективном обзоре «Математика в СССР за 30 лет» (Гостехиздат, 1948).

ГЛАВА I

ПРОИЗВОДНЫЕ

§ 1. Производная и дифференциал

1. Задачи, приводящие к понятию производной. Самым важным понятием дифференциального исчисления является понятие производной. Рассмотрим несколько задач конкретного характера, приводящих к этому понятию.

А. ЗАДАЧА О СКОРОСТИ. Пусть точка M движется по некоторой прямой [1]). Как известно, *средней скоростью* этой точки за какой-нибудь промежуток времени .называется отношение расстояния, пройденного точкой за этот промежуток времени, к его продолжительности.

Легко видеть, что знание средней скорости точки за тот или иной промежуток времени не даёт нам представления о характере движения в отдельные моменты этого промежутка. В связи с этим в механике приходится рассматривать также понятие *скорости точки в данный момент времени*. Под этим понимается предел средней скорости точки за бесконечно малый [2]) промежуток времени, начинающийся в данный момент (или оканчивающийся в этот момент).

Иногда, чтобы подчеркнуть отличие этого понятия от ранее определённой средней скорости, говорят об «истинной» скорости точки в данный момент.

Пусть на прямой, по которой движется точка, выбрана некоторая начальная точка O, выбрана определённая единица длины и прямой приписано определённое направление. Тогда положение движущейся точки M в каждый момент времени может быть определено указанием её расстояния OM от начальной точки O (рис. 1). Это расстояние будет некоторым (положительным, отрицательным или равным

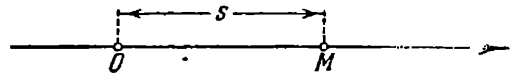


Рис. 1.

[1]) В случае не прямолинейного движения скорость имеет векторный характер и её определение более сложно.

[2]) То-есть такой, продолжительность которого стремится к нулю.

нулю) числом s, очевидно зависящим от того, о каком моменте времени идёт речь. С другой стороны, сам этот момент времени определяется указанием числа t единиц времени, отделяющих его от некоторого определённого начального момента.

Таким образом, $OM=s$ будет *функцией* аргумента t

$$s=f(t), \tag{1}$$

и знание этой функции полностью определяет движение точки. Равенство (1) называется *уравнением движения*. Вот примеры таких уравнений:

$$s=t^2, \quad s=2t^3+1, \quad s=\sin t, \quad s=\frac{gt^2}{2}.$$

Посмотрим, как найти «истинную» скорость точки M в момент t (т. е. в момент, отделённый от начального момента промежутком времени продолжительностью в t единиц), если известно уравнение движения (1).

Рассмотрим наряду с моментом t другой момент времени $t+\Delta t$. В этот момент движущаяся точка M находится на расстоянии

$$s+\Delta s=f(t+\Delta t)$$

от начальной точки O. Поэтому расстояние Δs, пройденное точкой за промежуток времени, начинающийся в момент t и кончающийся в момент $t+\Delta t$, равно

$$\Delta s=f(t+\Delta t)-f(t).$$

Поскольку продолжительность этого промежутка равна Δt, ясно, что средняя скорость за этот промежуток времени есть отношение

$$\frac{\Delta s}{\Delta t}=\frac{f(t+\Delta t)-f(t)}{\Delta t},$$

а тогда истинная скорость v точки M в момент t будет равна пределу этого отношения при стремящемся к нулю Δt:

$$v=\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t}=\lim_{\Delta t\to 0}\frac{f(t+\Delta t)-f(t)}{\Delta t}. \tag{I}$$

Таким образом, физическая задача нахождения скорости приводит к чисто аналитической задаче нахождения предела (I).

Рассмотрим другую задачу, также приводящую к подобному пределу.

Б. ЗАДАЧА О КАСАТЕЛЬНОЙ. В элементарной геометрии, где изучается одна лишь кривая линия — окружность, касательная к окружности определяется как прямая, которая имеет с этой окружностью всего лишь одну общую точку. Если рассматривать произвольные кривые, то такое определение касательной уже не будет удовлетво-

рительным. Вряд ли естественно считать, что ось Oy является касательной к параболе $y = x^2$ (рис. 2), хотя они и имеют всего лишь одну общую точку.

В связи с этим в науке принято другое, более общее, определение касательной.

Именно, *касательной* к кривой K в данной её точке M (которая называется *точкой касания*) называется прямая MT, являющаяся предельным положением [1]) секущей MS, проведённой через M и другую точку N кривой K, когда эта другая точка, оставаясь на K, стремится к совпадению с M (рис. 3).

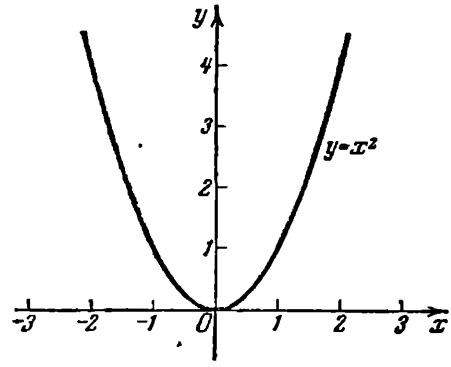


Рис. 2.

Легко убедиться, что в случае, когда K является окружностью, новое определение касательной равносильно даваемому в средней школе.

Поставим теперь задачу проведения касательной к кривой

$$y = f(x), \tag{2}$$

где $f(x)$ — некоторая непрерывная функция, если задана точка касания $M(x, y)$.

Поскольку нам известна точка M, то для определения касательной MT достаточно знать её угловой коэффициент k, т. е. тангенс угла α, под которым MT наклонена к оси Ox (рис. 4).

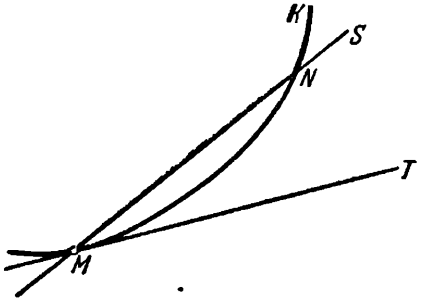


Рис. 3.

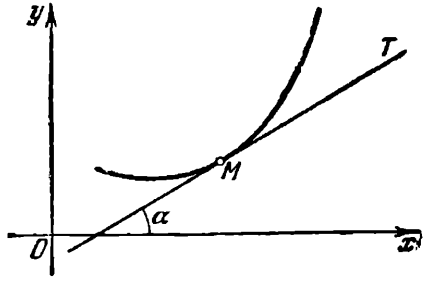


Рис. 4.

Чтобы определить этот угловой коэффициент, возьмём на нашей кривой ещё одну точку $N(x + \Delta x, y + \Delta y)$ и проведём через M и N секущую MN. Из аналитической геометрии известно, что угловой коэффициент секущей будет

$$k^* = \frac{\Delta y}{\Delta x}.$$

[1]) Это означает, что угол между прямыми MT и MS стремится к нулю.

Ввиду того, что касательная есть предельное положение секущей, ясно [1]), что

$$k = \lim_{N \to M} k^*. \tag{3}$$

Так как обе точки M и N лежат на кривой (2), то

$$y = f(x), \quad y + \Delta y = f(x + \Delta x),$$

откуда

$$\Delta y = f(x + \Delta x) - f(x).$$

С другой стороны, из факта стремления точки N к точке M вытекает, что Δx стремится к нулю [2]). Поэтому равенство (3) можно переписать и в такой форме

$$k = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \tag{II}$$

Таким образом, дело свелось к нахождению предела (II), который ничем, кроме обозначений, не отличается от предела (I).

Рассмотрим ещё один конкретный вопрос, который снова приведёт нас к тому же пределу.

в. ЗАДАЧА О ПЛОТНОСТИ. *Средней плотностью* прямолинейного стержня называется отношение его массы к его длине (мы имеем здесь в виду линейную плотность стержня; к этому понятию приходят, когда пренебрегают толщиной и шириной стержня). Чтобы получить более точное представление о характере распределения массы вдоль стержня, вводят понятие «истинной» плотности стержня в данной его точке. Под этим понимают предел средней плотности бесконечно малого участка стержня, стягивающегося в эту точку.

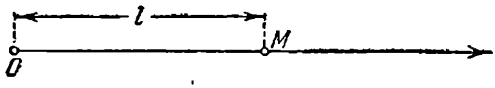


Рис. 5.

Будем характеризовать положение точки M на стержне её расстоянием $OM = l$ от одного из концов O стержня (рис. 5). Тогда, обозначив через m массу участка OM, мы, очевидно, получим, что m окажется функцией l

$$m = f(l), \tag{4}$$

причём знание зависимости (4) полностью определит нам, как распределяется масса стержня.

Поставим вопрос, как, зная уравнение (4), найти плотность стержня в данной его точке M, где $OM = l$.

---

[1]) Строгое доказательство равенства (3) без труда вытекает из непрерывности функции $\operatorname{tg} \alpha$.

[2]) В силу непрерывности $f(x)$ справедливо и обратное: при $\Delta x \to 0$ будет $N \to M$.

С этой целью рассмотрим[1]), кроме M, другую точку N стержня, для которой $ON=l+\Delta l$. Тогда масса участка ON будет

$$m+\Delta m=f(l+\Delta l),$$

откуда следует, что масса участка MN равна

$$\Delta m=f(l+\Delta l)-f(l).$$

Поэтому средняя плотность этого участка равна

$$\frac{\Delta m}{\Delta l}=\frac{f(l+\Delta l)-f(l)}{\Delta l}.$$

Истинная плотность p стержня в точке M есть предел этого отношения, когда N стремится к M, т. е. когда $\Delta l\to 0$.

Таким образом,

$$p=\lim_{\Delta l\to 0}\frac{\Delta m}{\Delta l}=\lim_{\Delta l\to 0}\frac{f(l+\Delta l)-f(l)}{\Delta l}. \qquad \text{(III)}$$

Мы снова приходим к необходимости рассмотрения того же предела, что и выше.

2. Определение производной. Рассмотрим теперь тот предел, к которому мы были приведены разобранными в п°1 конкретными задачами, не интересуясь уже его происхождением, а обращая внимание лишь на чисто математическую сторону дела.

Пусть на некотором открытом промежутке (a, b) задана непрерывная функция

$$y=f(x).$$

Проделаем следующие 5 операций:

1) Закрепим точку $x\in(a, b)$ и найдём соответствующее значение функции $y=f(x)$.

2) Придадим аргументу x отличное от нуля приращение Δx, не выводящее из промежутка (a, b), и найдём значение функции $y+\Delta y=f(x+\Delta x)$, соответствующее новому значению аргумента.

3) Вычислим приращение функции

$$\Delta y=f(x+\Delta x)-f(x).$$

4) Составим отношение

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

5) Устремим Δx к нулю и займёмся отысканием предела

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}. \qquad (5)$$

[1]) Читателю рекомендуется сделать чертёж.

Этот предел (если он существует[1])) и называется *производной* функции $f(x)$ в точке x.

Таким образом, *производная есть предел отношения приращения функции к вызвавшему его бесконечно малому приращению аргумента.*

Обозначается производная функции $y=f(x)$ через $f'(x)$, или через y'. Первое обозначение полнее, ибо в нём явно указана та точка x, в которой находится производная.

Действие нахождения производной какой-нибудь функции называется *дифференцированием* этой функции. Соответственно этому ту точку, в которой находится производная (т. е. то значение x аргумента, которое закрепляется в первой из вышеуказанных 5 операций), мы будем называть *точкой дифференцирования.*

Иллюстрируем данное определение несколькими примерами.

Пример 1. Найти производную функции $y=x^2$ в точке $x=5$.

Здесь $y=25$, $y+\Delta y=(5+\Delta x)^2=25+10\Delta x+(\Delta x)^2$. Значит,

$$\Delta y=10\Delta x+(\Delta x)^2 \quad \text{и} \quad \frac{\Delta y}{\Delta x}=10+\Delta x.$$

Устремляя Δx к нулю и находя соответствующий предел, получаем:

$$y'=10.$$

Пример 2. Найти производную функции $y=x^2$ в точках $x=4$, $x=1$, $x=0$, $x=12$.

Чтобы не повторять рассуждения для каждой из этих точек наново, проведём выкладку в общем виде, обозначив точку дифференцирования буквой x. Тогда при само собой понятных обозначениях мы получим:

$$y=x^2,\quad y+\Delta y=(x+\Delta x)^2=x^2+2x\Delta x+(\Delta x)^2,$$
$$\Delta y=2x\Delta x+(\Delta x)^2,$$
$$\frac{\Delta y}{\Delta x}=2x+\Delta x,$$
$$y'=2x. \qquad (6)$$

Подчеркнём ещё раз, что то значение x, которое фигурирует в правой части равенства (6), есть не что иное, как точка дифференцирования. Полагая, в частности, $x=4$, $x=1$, $x=0$, $x=12$, находим $y'=8$, $y'=2$, $y'=0$, $y'=24$.

Из рассмотренного примера видно преимущество дифференцирования функции в общем виде, т. е. при буквенном обозначении точки дифференцирования.

---

[1]) Мы имеем в виду конечный предел. Иногда вводятся «несобственные числа» $+\infty$ и $-\infty$ и тогда становится возможным говорить о бесконечных значениях производной, но мы этого делать не будем.

Пример 3. Найти производную функции $y=\sqrt{x}$ в точке x $(x>0)$. Не вдаваясь в пояснения, получаем:

$$y=\sqrt{x},\quad y+\Delta y=\sqrt{x+\Delta x},\quad \Delta y=\sqrt{x+\Delta x}-\sqrt{x},$$
$$\frac{\Delta y}{\Delta x}=\frac{\sqrt{x+\Delta x}-\sqrt{x}}{\Delta x}=\frac{1}{\sqrt{x+\Delta x}+\sqrt{x}},\quad y'=\frac{1}{2\sqrt{x}}.$$

Всякую функцию $y=f(x)$ можно изобразить графически. Сделав это, поставим вопрос о проведении касательной к полученной кривой (т. е. к графику функции). Этот вопрос мы рассматривали в п° 1. Сопоставляя полученное там выражение (II) с определением производной, получаем важное предложение:

Теорема. *Производная $y'=f'(x)$ геометрически представляет собой угловой коэффициент касательной к графику дифференцируемой функции, проведённой в точке, абсцисса которой есть точка дифференцирования.*

Общенаучное значение производной довольно отчётливо выступает в первой из рассмотренных выше задач. Действительно, в этой задаче выясняется, что в том случае, когда дифференцируемая функция представляет собой путь, пройденный движущейся точкой, а независимая переменная есть время, то производная представляет собой скорость движения.

Но в любом процессе, в котором приходится рассматривать две связанные между собой и изменяющиеся величины x и y, можно говорить о скорости изменения одной из них по отношению к другой. Нетрудно видеть, что точной характеристикой скорости изменения y по отношению к x как раз и будет производная y'.

Нет надобности говорить о том, насколько часто в науке приходится иметь дело с подобной скоростью изменения одной величины по отношению к другой. Например, если x есть время, а y — количество электричества, протекшего через сечение проводника за время x, то y' будет не что иное, как сила тока.

Легко увеличить число примеров такого рода.

3. Дифференцируемость и непрерывность. Односторонние производные. Производная выше была определена как предел некоторой переменной величины. Однако ведь не всякая переменная величина стремится к определённому пределу. В связи с этим и не всякая функция имеет производную. Легко показать, что для наличия производной у какой-либо функции необходима непрерывность этой функции. В самом деле, справедлива

Теорема. *Если у функции $y=f(x)$ в точке x существует производная* [1]*, то в этой точке функция непрерывна.*

---

[1]) В этом случае говорят, что функция $y=f(x)$ *дифференцируема* в точке x. (Выше мы употребляли выражение «дифференцируемая функция» в несколько ином смысле, понимая под ним функцию, которая подвергается процессу дифференцирования.)

Для доказательства заметим, что, давая аргументу (исходя из значения x) приращение $\Delta x \neq 0$ и обозначая через Δy соответствующее приращение функции, мы можем это последнее приращение записать так:

$$\Delta y = \frac{\Delta y}{\Delta x}\Delta x.$$

Если $\Delta x \to 0$, то отношение $\frac{\Delta y}{\Delta x}$ стремится к конечному [1]) пределу $y' = f'(x)$. Поэтому

$$\lim_{\Delta x \to 0} \Delta y = 0,$$

что и означает непрерывность функции $f(x)$ в точке x (т. е. при том значении аргумента, к которому прибавлялось Δx).

В связи с последней теоремой возникает вопрос, не будет ли условие непрерывности функции в какой-либо точке и достаточно для наличия производной в этой точке.

На этот вопрос приходится дать отрицательный ответ. Например, всюду непрерывная функция $y = |x|$ в точке $x = 0$ производной не имеет. В самом деле, здесь в зависимости от того, будет ли $\Delta x > 0$ или $\Delta x < 0$, окажется

$$\frac{\Delta y}{\Delta x} = \pm 1.$$

Поэтому в нашем случае предела

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

не существует. Это видно и геометрически, ибо график функции $y = |x|$ имеет вид ломаной, изображённой на рис. 6. Ясно, что в начале координат у этой ломаной нет касательной.

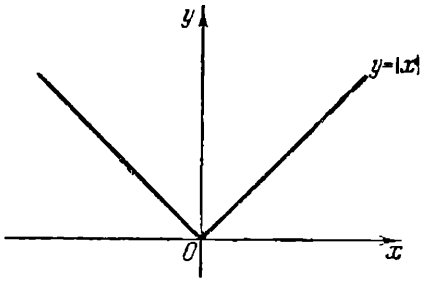


Рис. 6.

Рассмотренный пример очень прост, так как у функции $y = |x|$ в точке $x = 0$ существуют так называемые «правосторонняя» и «левосторонняя» производные. Правосторонней производной функции $f(x)$ в точке x называется предел выражения

$$\frac{f(x+\Delta x) - f(x)}{\Delta x},$$

в котором Δx стремится к нулю, принимая лишь положительные значения. Обозначается правосторонняя производная через $f'_+(x)$. Таким образом

$$f'_+(x) = \lim_{\substack{\Delta x \to 0 \\ \Delta x > 0}} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

[1]) См. сноску на стр. 308.

Аналогично левосторонняя производная $f'_-(x)$ определяется формулой

$$f'_-(x) = \lim_{\substack{\Delta x \to 0 \\ \Delta x < 0}} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

Для того чтобы $f(x)$ в точке x [1]) имела производную $f'(x)$, необходимо и достаточно, чтобы она имела в этой точке как правостороннюю, так и левостороннюю производные и чтобы эти последние производные были равны между собой.

Так как у функции $f(x) = |x|$ будет $f'_+(0) = +1$, $f'_-(0) = -1$, то этим и объясняется отсутствие у неё обыкновенной производной $f'(0)$.

Более интересен пример всюду непрерывной функции

$$f(x) = \begin{cases} x \sin \dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0), \end{cases}$$

у которой не существуют ни $f'_+(0)$, ни $f'_-(0)$.

В самом деле, если $\Delta x > 0$, то

$$\frac{f(0+\Delta x) - f(0)}{\Delta x} = \sin \frac{1}{\Delta x},$$

а эта величина не стремится ни к какому пределу при $\Delta x \to 0$, а колеблется между $+1$ и -1 бесконечное множество раз, принимая как эти, так и все промежуточные значения. Указанное явление отчётливо видно на рис. 7, где изображена часть графика

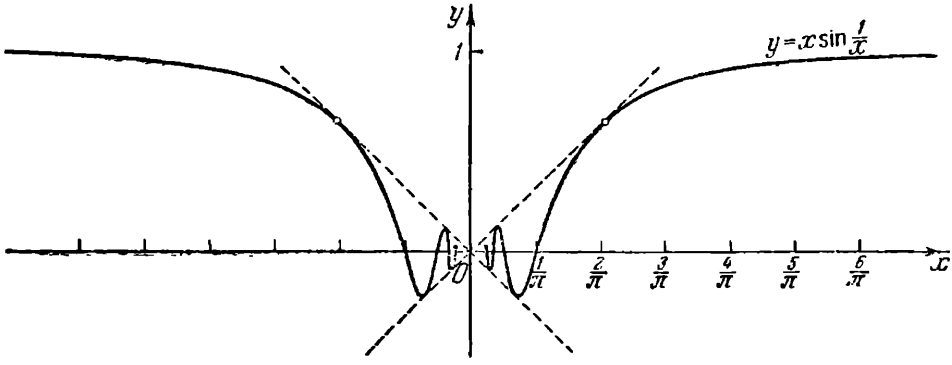


Рис. 7.

[1]) Здесь, как и выше, речь идёт о функции, заданной в открытом промежутке (a, b), содержащем точку x. Для функции $f(x)$, заданной в замкнутом отрезке $[a, b]$, также часто приходится рассматривать $f'_+(a)$ и $f'_-(b)$. Иногда для краткости не подчёркивают односторонности этих производных и говорят просто об $f'(a)$ и $f'(b)$.

функции $f(x)$. Ясно, что секущая, проведённая через начало координат и какую-либо точку $N(x, y)$ графика, будет по мере приближения N к началу совершать бесконечное множество колебаний между прямыми $y=+x$ и $y=-x$ и не будет стремиться ни к какому предельному положению.

В обоих рассмотренных примерах мы обнаружили отсутствие производной в отдельной точке. Такие примеры были известны и в XVIII в. Однако в ту эпоху существовало (явно не сформулированное) ошибочное убеждение, что непрерывная функция (это понятие, кстати сказать, тоже точно не определялось) имеет производную всюду, кроме разве лишь отдельных точек.

Ещё в 1834 г. великий русский математик Н. И. Лобачевский отчётливо различал свойства непрерывности и дифференцируемости функции. Тем не менее и в XIX в. были попытки доказать, что непрерывная функция, кроме как в отдельных точках, имеет производную. Этим попыткам положил конец К. Вейерштрасс, в 1871 г. построивший пример *непрерывной функции, которая нигде не имеет производной.* Но первый пример такого рода функции был построен гораздо раньше (не позднее 1830 г.) знаменитым чешским математиком Б. Больцано. Пример этот долгое время оставался неизвестным и был опубликован лишь в 20-х годах текущего века.

4. Производные простейших элементарных функций. В настоящем и ближайших двух пп° мы докажем, что *все элементарные функции имеют производные во всех точках своих областей задания, за исключением лишь отдельных особых точек, причём эти производные сами являются элементарными функциями точки дифференцирования.*

Целью настоящего п° является доказательство следующих формул [1])

1) $(C)'=0$,

2) $(x)'=1$,

3) $(x^n)'=nx^{n-1}$,

4) $(x^a)'=ax^{a-1}$,

5) $(\sin x)'=\cos x$,

6) $(\cos x)'=-\sin x$,

7) $(\operatorname{tg} x)'=\dfrac{1}{\cos^2 x}$,

8) $(\operatorname{ctg} x)'=-\dfrac{1}{\sin^2 x}$,

9) $(a^x)'=a^x \ln a$,

10) $(e^x)'=e^x$,

11) $(\ln x)'=\dfrac{1}{x}$,

12) $(\log_a x)'=\dfrac{1}{x \ln a}$.

Буква x, фигурирующая в правых частях этих формул, обозначает точку дифференцирования. В формуле 3) показатель n озна-

[1]) На функциях $\sec x$ и $\operatorname{cosec} x$ мы не останавливаемся, так как они вообще применяются редко. Обратные тригонометрические функции рассматриваются в п° 5.

чает натуральное число. Показатель же a, входящий в формулу 4), обозначает любое действительное число. Соответственно этому функция x^n рассматривается как заданная на всей оси, функция же x^a определена не везде. Именно, если a — иррациональное число, то эта функция определена только для $x>0$ (если $a>0$, то и для $x=0$); если же a рационально и не равно 0, то, записывая его в виде несократимой дроби $\frac{p}{q}$ $(q\geqslant 1)$, найдём, что при q нечётном функция x^a определена как для $x>0$, так и для $x<0$ (если $p>0$, то и для $x=0$), а при q чётном — только для $x>0$ (если $p>0$, то и для $x=0$). Точно так же и некоторые другие из функций, приведённых в нашей таблице, теряют смысл при тех или иных значениях аргумента. Например, функция $\operatorname{tg} x$ теряет смысл при $x=\frac{\pi}{2}$, функция $\ln x$ — при $x\leqslant 0$ и т. п. Однако, как это будет ясно из ниже приводимых доказательств, *каждая из формул* 1) — 12) *справедлива при любом* x, *при котором имеют смысл обе части этой формулы.*

При доказательстве нам придётся использовать следующие формулы теории пределов:

$$\lim_{z\to 0}\frac{\ln(1+z)}{z}=1, \tag{А}$$

$$\lim_{z\to 0}\frac{(1+z)^a-1}{z}=a, \tag{Б}$$

$$\lim_{z\to 0}\frac{a^z-1}{z}=\ln a, \tag{В}$$

$$\lim_{z\to 0}\frac{\sin z}{z}=1. \tag{Г}$$

Формула (Г) выведена в предыдущей статье (см. стр. 192). Дадим краткие выводы формул (А), (Б), (В).

Как известно (см. стр. 200), число e может быть определено как предел выражения

$$\left(1+\frac{1}{x}\right)^x$$

при $x\to\pm\infty$. Если положить

$$x=\frac{1}{z},$$

то мы приходим к равенству

$$\lim_{z\to 0}(1+z)^{\frac{1}{z}}=e,$$

Отсюда на основании непрерывности логарифмической функции получается

$$\lim_{z\to 0}\frac{\ln(1+z)}{z}=\lim_{z\to 0}\ln(1+z)^{\frac{1}{z}}=\ln e=1,$$

чем и доказана формула (А).

Переходя к выводу формулы (Б), положим

$$(1+z)^a-1=u.$$

Ясно, что при $z\to 0$ будет и $u\to 0$. При этом

$$(1+z)^a=1+u$$

и

$$a\ln(1+z)=\ln(1+u).$$

Таким образом,

$$\frac{(1+z)^a-1}{z}=\frac{u}{z}=\frac{u}{\ln(1+u)}\cdot\frac{a\ln(1+z)}{z}.$$

Отсюда и из (А) вытекает (Б).

Полагая, наконец,

$$a^z-1=u,$$

мы видим, что $u\to 0$ при $z\to 0$. С другой стороны,

$$z=\frac{\ln(1+u)}{\ln a}$$

и потому

$$\frac{a^z-1}{z}=\frac{u}{\ln(1+u)}\ln a,$$

откуда вытекает (В).

Переходя к доказательствам самих формул 1) — 12), заметим, что формулы 1) и 2) очевидны. В самом деле, если $y=C$, т. е. y есть постоянная величина, то приращение Δy этой функции, вызванное приращением Δx аргумента, равно нулю, откуда и следует, что $(C)'=0$. Точно так же, если $y=x$, то и $\Delta y=\Delta x$, откуда сразу вытекает формула 2).

Установим формулу 3). Если $y=x^n$, то

$$y+\Delta y=(x+\Delta x)^n,\quad \Delta y=(x+\Delta x)^n-x^n$$

и, следовательно,

$$\frac{\Delta y}{\Delta x}=\frac{(x+\Delta x)^n-x^n}{(x+\Delta x)-x}=(x+\Delta x)^{n-1}+(x+\Delta x)^{n-2}x+\ldots+x^{n-1}.$$

Отсюда при Δx, стремящемся к нулю, получаем:

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\underbrace{x^{n-1}+x^{n-1}+\ldots+x^{n-1}}_{n\text{ раз}}=nx^{n-1}.$$

Формула 3) установлена.

Переходя к формуле 4), допустим, что нами рассматривается такое отличное от нуля значение аргумента x, при котором функция x^a определена. Тогда она определена и для близких значений $x+\Delta x$ и

$$y+\Delta y=(x+\Delta x)^a=x^a\left(1+\frac{\Delta x}{x}\right)^a.$$

Значит,

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}x^a\cdot\frac{\left(1+\frac{\Delta x}{x}\right)^a-1}{\Delta x}=x^{a-1}\lim_{\Delta x\to 0}\frac{\left(1+\frac{\Delta x}{x}\right)^a-1}{\frac{\Delta x}{x}}.$$

Положив $\frac{\Delta x}{x}=z$ и применив формулу (Б), мы и получаем 4).

Приведённое доказательство не годится для $x=0$. Проведём рассуждение для этого случая отдельно. Согласно сделанной выше оговорке, мы должны предполагать, что при $x=0$ имеют смысл обе части формулы 4), т. е. что[1]) $a>1$. Но тогда при $x=0$ будет $y=x^a=0$, а $y+\Delta y=(\Delta x)^a$.

Значит,

$$\frac{\Delta y}{\Delta x}=(\Delta x)^{a-1}$$

и, стало быть, $y'=0$, что согласуется с формулой 4) при $a>1$ и $x=0$. (Полезно отметить, что при $a=\frac{2m-1}{2n}$, где m и n — натуральные числа, функция x^a определена лишь для $x\geqslant 0$. Соответственно этому, формула 4) доставляет нам при таком a и при $x=0$ п р а в о с т о р о н н ю ю производную.)

Формулы 5) и 6) устанавливаются совершенно аналогично. Остановимся для примера на выводе формулы 5).

Имеем:

$$y=\sin x,\quad y+\Delta y=\sin(x+\Delta x).$$

Стало быть,

$$\Delta y=\sin(x+\Delta x)-\sin x=2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right)$$

и

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\left(\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\right)\cdot\lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right).$$

Первый из сомножителей правой части согласно (Г) равен 1, а второй (в силу непрерывности функции $\cos x$) равен $\cos x$. Значит, $y'=\cos x$, и формула 5) доказана.

---

[1]) Напомним, что символ 0^0 лишён смысла.

Формулы 7) и 8) также доказываются аналогично, и мы остановимся лишь на первой из них. Предполагая, что $x \neq (2n-1)\frac{\pi}{2}$, имеем:

$$y = \operatorname{tg} x, \quad y + \Delta y = \operatorname{tg}(x + \Delta x).$$

Отсюда [1])

$$\Delta y = \frac{\sin \Delta x}{\cos x \cdot \cos(x + \Delta x)}$$

и остальное ясно.

Формула 9) устанавливается так:

$$y = a^x, \quad y + \Delta y = a^{x+\Delta x}, \quad \Delta y = a^x (a^{\Delta x} - 1),$$

и остаётся сослаться на формулу (B).

Формула 10) есть частный случай 9).

Переходя к формуле 11), допустим, что $x > 0$. Тогда

$$y = \ln x, \quad y + \Delta y = \ln(x + \Delta x),$$

$$\Delta y = \ln(x + \Delta x) - \ln x = \ln\left(1 + \frac{\Delta x}{x}\right).$$

Значит, на основании (A) получаем:

$$y' = \lim_{\Delta x \to 0} \frac{\ln\left(1 + \frac{\Delta x}{x}\right)}{\Delta x} = \frac{1}{x} \lim_{\Delta x \to 0} \frac{\ln\left(1 + \frac{\Delta x}{x}\right)}{\frac{\Delta x}{x}} = \frac{1}{x}.$$

Наконец, для $y = \log_a x$ имеем $a^y = x$, откуда $y = \frac{\ln x}{\ln a}$. Повторив для $\frac{\ln x}{\ln a}$ рассуждение, которое мы только что провели для $\ln x$, мы придём к 12).

В заключение остановимся на вопросе, почему в формулах математического анализа постоянно применяются не десятичные, а натуральные логарифмы.

На первый взгляд этот факт представляется весьма странным, ибо десятичные логарифмы обладают важными преимуществами, делающими их незаменимым пособием при вычислениях. Напомним эти преимущества. Из самого определения логарифма вытекает, что

$$\lg 1 = 0, \quad \lg 10 = 1, \quad \lg 100 = 2, \quad \lg 1000 = 3, \ldots$$

Если требуется найти логарифм числа, не являющегося целой степенью 10, например числа 637, то рассуждают так. Число 637

---

[1]) Как известно,

$$\operatorname{tg} A - \operatorname{tg} B = \frac{\sin(A - B)}{\cos A \cos B},$$

лежит между 100 и 1000. Отсюда на основании монотонности логарифма вытекает, что

$$\lg 100 < \lg 637 < \lg 1000,$$

или

$$2 < \lg 637 < 3.$$

Значит,

$$\lg 637 = 2, \ldots,$$

т. е. целая часть логарифма (его «характеристика») определяется с одного взгляда на число, и дело сводится к нахождению его дробной части («мантиссы»).

Далее, числа 637 и 63,7 связаны формулой:

$$637 = 63{,}7 \cdot 10.$$

Значит,

$$\lg 637 = \lg 63{,}7 + \lg 10 = \lg 63{,}7 + 1.$$

Поэтому мантиссы логарифмов чисел 637 и 63,7 одинаковы. Таким образом, упомянутые преимущества десятичных логарифмов состоят в простоте нахождения характеристики и неизменности мантиссы при умножении числа на 10, или на целую степень 10.

Нетрудно понять, что источником этих преимуществ является то, что мы пользуемся десятичной системой счисления. Если бы мы пользовались двенадцатиричной системой счисления, то упомянутые преимущества принадлежали бы уже не десятичным логарифмам, а логарифмам с основанием 12. Применение в качестве основания системы счисления числа 10 объясняется вовсе не какими-либо особыми арифметическими достоинствами этого числа, а тем чисто физиологическим обстоятельством, что на руках у человека 10 пальцев. Считая предметы и загибая при этом пальцы, человеку на ранних ступенях культуры приходилось, дойдя до 10, делать какую-либо пометку: зарубку на дереве, или что-либо подобное. Эта пометка и являлась «счётной единицей второго разряда». Мы видим, что она оказывалась равной 10.

Таким образом, выбор десятичной системы логарифмов никакими теоретическими соображениями не обусловлен и потому никаких *теоретических* выгод и не доставляет. Это, конечно, ни в какой степени не понижает чисто вычислительных удобств этой системы и потому во всех численных расчётах всегда пользуются именно ею.

Однако в математике значительно бо́льшую роль играют не *численные*, а *буквенные* расчёты. При таких расчётах отмеченные выше преимущества десятичной системы логарифмов исчезают. В то же время возникает вопрос о выборе такого основания системы логарифмов, чтобы применяемые *буквенные формулы* имели по возможности более простой вид. И вот с этой-то новой

точки зрения натуральные логарифмы и оказываются наилучшими, ибо производная натурального логарифма

$$(\ln x)' = \frac{1}{x}$$

выражается более простой формулой, чем производная всякого другого логарифма

$$(\log_a x)' = \frac{1}{x \ln a}.$$

В частности, так как $\ln 10 = 2{,}30259\ldots$, то производная десятичного логарифма

$$(\lg x)' = \frac{1}{x \cdot 2{,}30259\ldots} = \frac{0{,}43429\ldots}{x}$$

выражается сложнее, чем логарифма натурального.

5. Дифференцирование обратных функций. Пусть функция $y = f(x)$ задана, непрерывна и строго монотонна в замкнутом промежутке $[a, b]$. Как известно (см. стр. 245), в этом случае существует функция, *обратная* для функции $f(x)$, т. е. такая однозначная функция $x = g(y)$, которая задана в замкнутом промежутке, содержащемся между $f(a)$ и $f(b)$, и сопоставляющая с каждым y из этого промежутка то (единственное) значение x из $[a, b]$, для которого $f(x) = y$. Эта обратная функция также непрерывна и строго монотонна [1]).

В этих условиях (и в этих обозначениях) справедлива

Теорема. *Если у функции $f(x)$ в точке x_0 существует отличная от нуля производная $f'(x_0)$, то у обратной функции $g(y)$ в соответствующей точке $y_0 = f(x_0)$ существует производная $g'(y_0)$, причём*

$$g'(y_0) = \frac{1}{f'(x_0)}.$$

Доказательство. Пусть $\Delta y \neq 0$ и точка $y_0 + \Delta y$ содержится между $f(a)$ и $f(b)$. Тогда, полагая

$$g(y_0 + \Delta y) = x_0 + \Delta x,$$

мы будем иметь, что $\Delta x \neq 0$ и $x_0 + \Delta x \in [a, b]$, причём

$$f(x_0 + \Delta x) = y_0 + \Delta y.$$

---

[1]) Если прямая функция $y = f(x)$ задана в открытом промежутке (a, b), то и обратная функция будет определена в открытом же промежутке, лежащем между $f(a)$ и $f(b)$. Один (или оба) из упомянутых в этой формулировке промежутков может быть и бесконечен [например, если $y = \operatorname{tg} x$, $-\frac{\pi}{2} < x < \frac{\pi}{2}$, то обратная функция $x = \operatorname{arctg} y$ определена в $(-\infty, +\infty)$].

Очевидно,

$$\frac{g(y_0+\Delta y)-g(y_0)}{\Delta y}=\frac{\Delta x}{f(x_0+\Delta x)-f(x_0)}.$$

Устремляя Δy к нулю, мы получим, что и $\Delta x \to 0$ (ибо функция $x=g(y)$ непрерывна). Стало быть, правая часть предыдущего равенства стремится (при $\Delta y \to 0$) к $\frac{1}{f'(x_0)}$. Теорема доказана.

Краткая, но выразительная запись этой теоремы такова:

$$y'_x=\frac{1}{x'_y}.$$

Применим теорему к доказательству формулы

$$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}} \qquad (-1<x<1).$$

Для этого рассмотрим функцию

$$y=\sin x,$$

причём будем изменять x только в замкнутом промежутке $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Здесь наша функция строго возрастает от -1 до $+1$. Кроме того, она непрерывна. Значит, она имеет обратную функцию

$$x=\arcsin y,$$

определённую в $[-1, +1]$ и сопоставляющую с каждым $y \in [-1, +1]$ то единственное значение $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, для которого $\sin x = y$.

Замечая, что производная $y'=\cos x$ отлична от нуля всюду, кроме точек $\pm\frac{\pi}{2}$, мы заключаем, что обратная функция $x=\arcsin y$ имеет производную при всех y, кроме $y=\pm 1$, причём

$$(\arcsin y)'=\frac{1}{\cos x},$$

где $x=\arcsin y$.

Так как $\cos^2 x+\sin^2 x=1$, то

$$\cos x=\pm\sqrt{1-\sin^2 x}=\pm\sqrt{1-y^2}.$$

Из двух знаков, стоящих перед радикалом, мы обязаны остановиться на знаке $+$, ибо $-\frac{\pi}{2}<x<\frac{\pi}{2}$ (ведь $x=\arcsin y$, а это значение лежит между $-\frac{\pi}{2}$ и $\frac{\pi}{2}$), а для таких x будет $\cos x>0$.

Итак,

$$(\arcsin y)'=\frac{1}{\sqrt{1-y^2}},$$

что лишь обозначением отличается от доказываемой формулы.

Совершенно аналогично устанавливаются еще 3 формулы:

$$(\arccos x)' = \frac{-1}{\sqrt{1-x^2}} \qquad (-1 < x < +1),$$

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2} \qquad (-\infty < x < +\infty),$$

$$(\operatorname{arcctg} x)' = \frac{-1}{1+x^2} \qquad (-\infty < x < +\infty).$$

6. Правила комбинирования формул дифференцирования. Формулы, выведенные в пп° 4 и 5, доказывают дифференцируемость ряда простейших элементарных функций и дают выражения для их производных. Однако на практике приходится иметь дело обычно не с самими рассмотренными простейшими функциями, а с теми или иными их комбинациями, например с результатами арифметических действий над этими функциями. Обратимся к вопросу дифференцирования подобных комбинаций простейших функций и прежде всего остановимся на задаче дифференцирования суммы, разности, произведения и частного таких функций.

Ясно, что здесь мы имеем дело с частным случаем следующей более общей задачи: пусть $u = u(x)$ и $v = v(x)$ — две (не обязательно элементарные) функции, заданные в одном и том же промежутке (a, b) и имеющие в некоторой точке x этого промежутка производные u' и v'. Требуется установить дифференцируемость в точке x каждой из функций [1])

$$u + v,\ u - v,\ uv,\ \frac{u}{v} \tag{7}$$

и выразить их производные через u' и v'.

Решается эта задача однотипным рассуждением, которое в качестве примера проведём для случая произведения uv.

Итак, пусть значения функций $u(x)$ и $v(x)$, отвечающие точке дифференцирования x, суть u и v. Придавая x приращение Δx [отличное от нуля и не выводящее из промежутка (a, b)], мы изменим значения $u(x)$ и $v(x)$. Пусть их новые значения суть $u + \Delta u$ и $v + \Delta v$. Тогда новое значение произведения $y = u(x)\,v(x)$ будет равно

$$y + \Delta y = (u + \Delta u)(v + \Delta v).$$

Вычитая отсюда старое значение $y = uv$ этого произведения и деля результат на Δx, находим

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} v + u \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x} \cdot \Delta v.$$

---

[1]) При рассмотрении частного $\frac{u}{v}$ предполагается, что в точке x знаменатель v отличен от нуля.

Устремив Δx к нулю, мы будем иметь по самому определению производной

$$\lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} = u', \quad \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} = v'.$$

Кроме того, и это весьма важно,

$$\lim_{\Delta x \to 0} \Delta v = 0.$$

В самом деле, раз функция $v(x)$ в точке x имеет производную, то она в этой точке непрерывна и потому стремление к нулю приращения аргумента влечёт за собой и стремление к нулю приращения функции.

Сопоставляя всё сказанное, мы видим, что

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = u'v + uv',$$

или, что то же самое,

$$(uv)' = u'v + uv'.$$

Аналогичные рассуждения приводят к соотношениям

$$(u + v)' = u' + v', \quad (u - v)' = u' - v', \quad \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2},$$

причём попутно устанавливается и самое существование производных у всех функций (7).

Заслуживает быть отмеченным тот частный случай дифференцирования произведения, когда один из сомножителей постоянен. В этом случае вышеприведённая формула принимает вид

$$(cu)' = cu',$$

т. е. *постоянный множитель можно выносить из-под знака производной.*

Теперь мы в состоянии находить производные довольно разнообразных комбинаций простейших элементарных функций, как, например [1]),

$$y = \frac{\sin x}{\sqrt{x}} e^x,$$

или

$$y = (2x^3 + 5x^2 - 7x + 2) \frac{\operatorname{arctg} x}{\ln x}$$

и т. п.

---

[1]) Обе эти функции дифференцируются в конце п° (примеры 2 и 3).

Однако мы не можем ещё найти производную функции

$$y = e^{\sin x},$$

ибо она не есть результат арифметических операций над простейшими элементарными функциями.

Чтобы иметь возможность дифференцировать функции подобного вида, нам понадобится ещё одно — и, пожалуй, важнейшее из всех — правило дифференцирования. Это — *правило дифференцирования сложных функций.*

Постановка вопроса здесь такова. Пусть

$$y = f(z)$$

есть функция, заданная в промежутке (A, B) и имеющая в точке z_0 этого промежутка производную

$$y'_z = f'(z_0).$$

Пусть, далее,

$$z = \varphi(x)$$

есть функция, заданная в промежутке (a, b) и обладающая тем свойством, что её значения удовлетворяют неравенству

$$A < \varphi(x) < B.$$

Тогда можно образовать функцию

$$f[\varphi(x)],$$

которая и называется *сложной* функцией аргумента x. Предположим, что точка $x_0 \in (a, b)$ переводится функцией $\varphi(x)$ в вышеупомянутую точку z_0, т. е. что $\varphi(x_0) = z_0$, и допустим, что существует производная $\varphi'(x_0) = z'_x$. Поставим вопрос о дифференцировании функции $f[\varphi(x)]$ в точке x_0.

Для решения этого вопроса нам понадобится одна вспомогательная формула, имеющая, впрочем, и самостоятельный интерес. Пусть $y = f(x)$ есть функция, имеющая в некоторой точке x производную $y'_x = f'(x)$. Придадим x приращение Δx, отличное от нуля, но не выводящее из промежутка задания функции, и обозначим через Δy соответствующее приращение функции. Так как отношение $\frac{\Delta y}{\Delta x}$ при стремлении Δx к нулю стремится к производной y'_x, а разность между переменной, имеющей предел, и этим пределом есть величина бесконечно малая, то величина

$$\alpha = \frac{\Delta y}{\Delta x} - y'_x \tag{8}$$

стремится к нулю вместе с Δx. Перепишем равенство (8) в другой форме:

$$\Delta y = y'_x \Delta x + \alpha \Delta x. \tag{9}$$

Эга новая формула, так же как и формула (8), имеет смысл только при $\Delta x \neq 0$, ибо величина α, являющаяся по существу функцией от Δx, $\alpha = \alpha(\Delta x)$, для $\Delta x = 0$ не определена. С другой стороны, непосредственно ясно, что при $\Delta x = 0$ будет и $\Delta y = f(x + \Delta x) - f(x) = 0$. Значит, достаточно произвольным образом доопределить функцию $\alpha(\Delta x)$ для $\Delta x = 0$, чтобы формула (9) стала верной и для этого Δx.

Положим

$$\alpha(\Delta x) = \begin{cases} \dfrac{\Delta y}{\Delta x} - y'_x & \text{при } \Delta x \neq 0, \\ 0 & \text{при } \Delta x = 0. \end{cases} \tag{10}$$

Тогда, как уже сказано, формула (9) будет верна, независимо от того, обращается Δx в нуль или нет, и сверх того, что здесь является основным, при любой последовательности $\{h_n\}$, стремящейся к нулю, будет

$$\lim \alpha(h_n) = 0,$$

опять-таки независимо от того, имеются в этой последовательности члены, равные нулю, или нет.

Заметив это, вернёмся к вопросу о дифференцировании функции $y = f[\varphi(x)]$.

Сохраняя введённые выше обозначения, придадим аргументу x сначала значение x_0, а затем значение $x_0 + \Delta x$, где $\Delta x \neq 0$ и $x_0 + \Delta x \in (a, b)$. Соответственно этому аргумент $z = \varphi(x)$ будет иметь значения $z_0 = \varphi(x_0)$ и $z_0 + \Delta z = \varphi(x_0 + \Delta x)$, причём, однако, не исключено, что $\Delta z = 0$. Полагая $y_0 = f(z_0)$ и $y_0 + \Delta y = f(z_0 + \Delta z)$, мы будем иметь по сказанному выше:

$$\Delta y = y'_z \Delta z + \alpha \Delta z,$$

где α стремится к нулю, когда Δz пробегает последовательность значений, стремящихся к нулю. Это справедливо, независимо от того, имеются ли среди членов упомянутой последовательности нули. Разделив предыдущее равенство на Δx и переходя к пределу при $\Delta x \to 0$, получим:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = y'_z \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} + 0 \cdot \lim_{\Delta x \to 0} \cdot \frac{\Delta z}{\Delta x}.$$

Таким образом, мы и получаем то правило дифференцирования, о котором говорили выше:

$$y'_x = y'_z \cdot z'_x. \tag{11}$$

Можно записать это правило и так:

$$\{f[\varphi(x)]\}' = f'[\varphi(x)] \cdot \varphi'(x),$$

однако первая запись выразительнее. Если назвать z промежуточным аргументом, то формула (11) будет означать, что *производная*

сложной функции равна её производной по промежуточному аргументу, умноженной на производную промежуточного аргумента.

Дадим краткую сводку всех доказанных нами формул и правил дифференцирования:

$$
\left.\begin{array}{ll}
1) & (C)' = 0, \\
2) & (x)' = 1, \\
3) & (x^n)' = nx^{n-1}, \\
4) & (x^a)' = ax^{a-1}, \\
5) & (\sin x)' = \cos x, \\
6) & (\cos x)' = -\sin x, \\
7) & (\operatorname{tg} x)' = \dfrac{1}{\cos^2 x}, \\
8) & (\operatorname{ctg} x)' = -\dfrac{1}{\sin^2 x}, \\
9) & (a^x)' = a^x \ln a, \\
10) & (e^x)' = e^x, \\
11) & (\ln x)' = \dfrac{1}{x}, \\
12) & (\log_a x)' = \dfrac{1}{x \ln a}, \\
13) & (\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}, \\
14) & (\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}, \\
15) & (\operatorname{arctg} x)' = \dfrac{1}{1+x^2}, \\
16) & (\operatorname{arcctg} x)' = -\dfrac{1}{1+x^2}, \\
17) & (u+v)' = u' + v', \\
18) & (u-v)' = u' - v', \\
19) & (uv)' = u'v + uv', \\
20) & (cu)' = cu', \\
21) & \left(\dfrac{u}{v}\right)' = \dfrac{u'v - uv'}{v^2}, \\
22) & y'_x = \dfrac{1}{x'_y}, \\
23) & y'_x = y'_z \cdot z'_x.
\end{array}\right\} (*)
$$

Пользуясь этими формулами, мы можем найти производную любой элементарной функции, причём и сама эта производная оказывается элементарной функцией точки дифференцирования.

Иллюстрируем выведенные формулы рядом примеров.

Пример 1. Найти y', если

$$y = 7x^3 + 8x^2 - 9x + 11.$$

Пользуясь правилами 17), 18) и 20) таблицы (*), находим:

$$y' = (7x^3)' + (8x^2)' - (9x)' + (11)' =$$
$$= 7(x^3)' + 8(x^2)' - 9(x)' + (11)' = 21x^2 + 16x - 9.$$

Аналогичным образом, для любого многочлена

$$y = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n$$

находим:

$$y' = a_0nx^{n-1} + a_1(n-1)x^{n-2} + a_2(n-2)x^{n-3} + \ldots + a_{n-2}2x + a_{n-1}.$$

Пример 2. Найти y', если

$$y = \frac{\sin x}{\sqrt{x}} e^x.$$

Пользуясь правилами 19) и 21) таблицы (*) и формулами 5), 4) и 10) той же таблицы, находим:

$$y' = \frac{\cos x \cdot \sqrt{x} - \sin x \cdot \frac{1}{2} x^{-\frac{1}{2}}}{x} e^x + \frac{\sin x}{\sqrt{x}} e^x.$$

Пример 3. Найти y', если

$$y = (2x^3 + 5x^2 - 7x + 2) \frac{\operatorname{arctg} x}{\ln x}.$$

Имеем:

$$y' = (6x^2 + 10x - 7) \frac{\operatorname{arctg} x}{\ln x} + (2x^3 + 5x^2 - 7x + 2) \frac{\frac{\ln x}{1+x^2} - \frac{\operatorname{arctg} x}{x}}{\ln^2 x}.$$

Пример 4. Найти y', если

$$y = e^{\sin x}.$$

Здесь, в отличие от предыдущих примеров, где мы имели дело с результатами арифметических действий над простейшими элементарными функциями, мы встретились со сложной функцией. Полагая

$$y = e^z, \quad z = \sin x,$$

получаем:

$$y'_x = e^z \cos x = e^{\sin x} \cos x.$$

Пример 5. Найти y', если

$$y = \sqrt{\operatorname{arctg} x}.$$

Полагая $y = \sqrt{z}$, $z = \operatorname{arctg} x$, находим:

$$y' = \frac{1}{2} (\operatorname{arctg} x)^{-\frac{1}{2}} \cdot \frac{1}{1+x^2}.$$

Пример 6. Найти y', если

$$y = \operatorname{tg}^3 \ln x.$$

Здесь надо дважды применить правило дифференцирования сложной функции. Именно, полагая

$$y = z^3, \quad z = \operatorname{tg} \ln x,$$

получим:

$$y' = 3z^2 \cdot z' = 3\operatorname{tg}^2 \ln x \cdot z'.$$

Чтобы найти z', полагаем

$$z = \operatorname{tg} u, \quad u = \ln x,$$

откуда

$$z' = \frac{1}{\cos^2 u} \cdot \frac{1}{x} = \frac{1}{x \cos^2 \ln x}$$

и окончательно:

$$y' = \frac{3\operatorname{tg}^2 \ln x}{x \cos^2 \ln x}.$$

Обычно промежуточных аргументов не выписывают, производя их введение в уме. Это позволяет избежать употребления большого числа различных букв, обозначая каждый раз функцию, подлежащую дифференцированию, одной и той же буквой (которая, таким образом, в процессе решения одной задачи обозначает последовательно разные вещи). При этом полезно руководствоваться следующим правилом: *если подлежащая дифференцированию функция является результатом целого ряда действий над аргументом x, то за промежуточный аргумент z следует принять результат всех этих действий, кроме последнего.*

Например, если подлежащая дифференцированию функция есть

$$y = \operatorname{tg}^5 \sqrt[3]{\ln \arcsin x},$$

то следует положить

$$z = \operatorname{tg} \sqrt[3]{\ln \arcsin x}.$$

Тогда $y = z^5$ и

$$y' = 5\operatorname{tg}^4 \sqrt[3]{\ln \arcsin x} \cdot (\operatorname{tg} \sqrt[3]{\ln \arcsin x})'.$$

Полагая теперь

$$z = \sqrt[3]{\ln \arcsin x},$$

получим:

$$y' = 5 \operatorname{tg}^4 \sqrt[3]{\ln \arcsin x} \frac{1}{\cos^2 \sqrt[3]{\ln \arcsin x}} (\sqrt[3]{\ln \arcsin x})'.$$

Теперь надо обозначить через z функцию

$$\ln \arcsin x,$$

и это даёт:

$$y' = 5\operatorname{tg}^4 \sqrt[3]{\ln \arcsin x} \cdot \frac{1}{\cos^2 \sqrt[3]{\ln \arcsin x}} \cdot \frac{1}{3} (\ln \arcsin x)^{-\frac{2}{3}} (\ln \arcsin x)'.$$

Полагая, наконец,

$$z = \arcsin x,$$

находим окончательно:

$$y' = 5\operatorname{tg}^4 \sqrt[3]{\ln \arcsin x} \cdot \frac{1}{\cos^2 \sqrt[3]{\ln \arcsin x}} \times$$

$$\times \frac{1}{3} (\ln \arcsin x)^{-\frac{2}{3}} \frac{1}{\arcsin x} \cdot \frac{1}{\sqrt{1-x^2}}.$$

На практике производят все указанные замены в уме и сразу пишут результат дифференцирования.

Например, если $y = \sqrt{\operatorname{arctg} e^{\sin x}}$, то

$$y' = \frac{1}{2} (\operatorname{arctg} e^{\sin x})^{-\frac{1}{2}} \cdot \frac{1}{1 + e^{2 \sin x}} e^{\sin x} \cdot \cos x.$$

7. Дифференциал. С понятием производной тесно связано понятие дифференциала, от которого и происходит название дифференциального исчисления.

Чтобы выяснить сущность понятия дифференциала, рассмотрим функцию $y=f(x)$, заданную в открытом промежутке (a, b) и имеющую в некоторой точке x этого промежутка производную $y'=f'(x)$. Придадим аргументу приращение Δx, отличное от нуля и такое, что $x+\Delta x \in (a, b)$. Пусть соответствующее приращение функции будет Δy Как мы видели в предыдущем п°, справедлива формула

$$\Delta y = y'\Delta x + \alpha\Delta x,$$

где α стремится к нулю вместе с Δx.

Полагая

$$\alpha\Delta x = \rho,$$

мы видим, что при бесконечно малом Δx переменная ρ также есть бесконечно малая величина и притом стремящаяся к нулю быстрее, чем Δx в следующем точном смысле слова

$$\lim_{\Delta x \to 0} \frac{\rho}{\Delta x} = 0.$$

Вообще, если мы имеем две бесконечно малые величины λ и μ, связанные между собой так, что

$$\lim \frac{\lambda}{\mu} = 0,$$

то говорят, что λ есть бесконечно малая более высокого порядка, чем μ. Таким образом, величина ρ есть бесконечно малая более высокого порядка, чем Δx. Грубо говоря, это означает, что при весьма малых Δx величина ρ будет ещё во много раз меньше, чем Δx. Поэтому при малых Δx величиной $\rho = \alpha\Delta x$ часто бывает возможно пренебречь, что приводит к приближённой формуле

$$\Delta y \approx y'\Delta x.$$

Эта формула показывает, что с точностью до малой высшего (сравнительно с Δx) порядка приращение функции Δy оказывается прямо пропорциональным приращению аргумента Δx. При этом коэффициентом пропорциональности здесь служит производная y' (вычисленная при том значении аргумента x, к которому прибавляется приращение Δx).

Произведение $y'\Delta x$ (которое лишь приближённо равно приращению Δy) и называется дифференциалом функции y в точке x, соответствующим приращению аргумента Δx. Таким образом, точное определение дифференциала таково: *дифференциалом функции* $y=f(x)$ *в точке* x, *соответствующим приращению* Δx, *называется*

произведение производной $y'=f(x)$, *вычисленной в упомянутой точке, на* Δx. Обозначается дифференциал функции $y=f(x)$ через dy или $df(x)$:

$$dy=y'\Delta x, \quad df(x)=f'(x)\,\Delta x. \tag{12}$$

Дифференциал dy является, таким образом, функцией двух не зависящих друг от друга [1]) аргументов — точки дифференцирования x и приращения Δx. В обозначении $df(x)$ указана только точка x, в обозначении же dy не указаны ни x, ни Δx. Поэтому оба обозначения недостаточно полны.

Дифференциал важен для науки потому, что (при закреплённой точке дифференцирования) он обладает двумя свойствами: 1) он прямо пропорционален приращению аргумента и 2) разность между ним и приращением функции (вызванным упомянутым приращением аргумента) есть бесконечно малая высшего порядка (сравнительно с Δx).

Первое из этих свойств обеспечивает простоту вычисления дифференциала, а второе позволяет при малых Δx приближённо заменять Δy на dy.

Иллюстрируем всё сказанное следующим простым примером. Пусть

$$y=2x^3+3x^2+4x+2.$$

Найти Δy и dy, если $x=2$, $\Delta x=0{,}001$.

Здесь $y=38$, $y+\Delta y=38{,}040015002$. Стало быть,

$$\Delta y=0{,}040015002.$$

С другой стороны, $y'=6x^2+6x+4$, а при $x=2$ будет $y'=40$ и потому

$$dy=0{,}04.$$

Таким образом, сохраняя предшествующие обозначения, находим

$$\rho=0{,}000015002.$$

Значит, относительная ошибка, которую мы делаем, заменяя Δy через dy, будет

$$\delta=\frac{\rho}{\Delta y}=\frac{0{,}000015002}{0{,}040015002}<\frac{0{,}000016}{0{,}04}=0{,}0004=0{,}04\%.$$

Ясно, что точность приближения при такой замене очень хороша.

---

[1]) В начале этого п° мы подчинили Δx условию $x+\Delta x\in(a, b)$. Тем самым мы ограничили выбор значения Δx в зависимости от значения x. Так как это представляет известные неудобства, то впредь мы отказываемся от указанного ограничения и определяем dy формулой $dy=y'\Delta x$ для совершенно произвольных значений Δx.

Вычислим ещё Δy и dy для той же функции $y = 2x^3 + 3x^2 + 4x + 2$ в той же исходной точке $x = 2$, но для $\Delta x = 3$. Здесь оказывается $y = 38$, $y + \Delta y = 347$ и потому

$$\Delta y = 309.$$

С другой же стороны, как и выше, $y' = 40$, так что

$$dy = 120.$$

Мы видим, что, в отличие от предыдущего примера, Δy и dy вовсе не близки друг к другу. Столь разительное различие двух приведённых примеров объясняется тем, что в первом из них величина Δx была весьма мала, а во втором нет. Вообще говоря, употребление приближённых формул без оценки совершаемой ошибки есть вещь довольно безответственная. Можно, например, говорить о том, что «7 приближённо равно 32», и хотя практическая бесполезность такого утверждения совершенно очевидна, но всё же формально оно правильно. Поэтому, говоря, что «при малых Δx величины Δy и dy приближённо равны», мы по существу не говорим ещё ничего определённого. Ниже будет показано, что для широкого класса функций совершаемая при замене Δy на dy ошибка не превосходит величины

$$M(\Delta x)^2,$$

где M есть верхняя граница абсолютной величины второй [1]) производной $f''(x)$ в рассматриваемом промежутке. Пока же мы будем довольствоваться указанными несколько расплывчатыми формулировками.

Для более полного уяснения понятия дифференциала установим его геометрический смысл. С этой целью построим график функции $y = f(x)$ и проведём к нему касательную MT в точке $M(x, y)$.

Пусть $N(x + \Delta x,\ y + \Delta y)$, где $y + \Delta y = f(x + \Delta x)$, есть другая точка этого графика (рис. 8). Из чертежа видно, что приращение функции Δy равно отрезку KN. В то же время отрезок KL равен

$$KL = MK \cdot \operatorname{tg}\alpha = y'\Delta x.$$

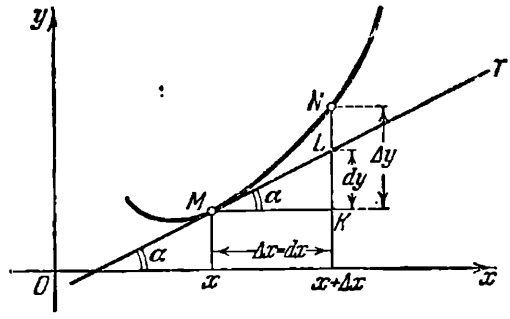


Рис. 8.

Иначе говоря, этот отрезок и есть дифференциал dy. Таким образом, *дифференциал геометрически представляет собой приращение ординаты точки, движущейся по касательной к графику функции*. Поскольку вблизи точки касания

[1]) См. стр. 334.

касательная весьма тесно примыкает к кривой, ясно, что для малых Δx замена Δy на dy не приводит к заметной погрешности.

Приведём несколько примеров нахождения дифференциалов.

1) Найти дифференциал функции $y=x^3$, если $x=2$, $\Delta x=0{,}1$. Здесь $y'=3x^2$, а так как точка дифференцирования есть $x=2$, то $y'=12$, откуда $dy=1{,}2$.

Мы видим, что согласно определению дифференциала его величину можно найти, только задав (кроме функции) точку дифференцирования x и приращение Δx. Однако в дальнейшем мы обе последние величины задавать численно не будем и сохраним для них буквенные обозначения.

2) Найти dy, если $y=e^x$. Здесь $y'=e^x$ и $dy=e^x\Delta x$. Ещё короче этот результат можно записать так:

$$de^x=e^x\Delta x.$$

Аналогично этому

3) $d\operatorname{tg}x=\dfrac{\Delta x}{\cos^2 x}$, 4) $d\ln x=\dfrac{\Delta x}{x}$.

Точно так же

5) $$d\operatorname{tg}^6\sqrt[3]{\ln\sin x}=$$
$$=6\operatorname{tg}^5\sqrt[3]{\ln\sin x}\,\frac{1}{\cos^2\sqrt[3]{\ln\sin x}}\cdot\frac{1}{3}(\ln\sin x)^{-\frac{1}{3}}\frac{1}{\sin x}\cos x\cdot\Delta x.$$

В последнем примере очень наглядно выступает, что нахождение дифференциала по существу сводится к нахождению производной (а значит, и наоборот). Именно поэтому действие нахождения производной и называется «дифференцированием».

Очень поучителен пример нахождения дифференциала функции $y=x$. Именно, здесь $y'=1$ и поэтому $dy=\Delta x$, или

$$dx=\Delta x.$$

Таким образом, *дифференциал независимой переменной совпадает с её приращением* [1]).

Отсюда следует, что формулу (12)

$$dy=y'\Delta x,$$

дающую самое определение дифференциала, можно переписать и так:

$$dy=y'dx. \qquad (13)$$

---

[1]) Легко видеть, что это верно для всякой линейной функции $y=ax+b$.

В свою очередь последняя формула приводит к возможности обозначать производную y' символом

$$\frac{dy}{dx}.$$

В книгах технического характера последнее обозначение применяется чаще, чем y'.

Формула (13) обладает весьма важным свойством. Именно, *она остаётся справедливой и в том случае, когда x уже не служит независимой переменной, а является функцией другого аргумента t.* В самом деле, в этом случае y будет также (сложной) функцией t, и дифференциал dy, отвечающий приращению Δt, надо вычислять по формуле

$$dy = y'_t \Delta t.$$

Согласно правилу дифференцирования сложной функции

$$y'_t = y'_x x'_t,$$

откуда

$$dy = y'_x x'_t \Delta t.$$

Замечая, что $x'_t \Delta t = dx$, мы возвращаемся к формуле (13). Чтобы оценить это свойство формулы (13), следует сопоставить его с тем фактом, что формула (12) подобным свойством заведомо не обладает.

Действительно, в том случае, когда x есть функция от t, символы Δx и dx означают разные понятия и потому справедливость одной из формул (12) или (13) исключает справедливость другой. В отмеченном свойстве состоит существенное преимущество записи дифференциала dy в инвариантной форме (13). Например, если $y = x^3$, то равенство

$$dy = 3x^2 dx$$

справедливо, независимо от того, является ли x независимой переменной или функцией.

Приведём несколько примеров применения дифференциала.

Пример 1. Вычислить $(1+h)^a$, если h весьма мало. Чтобы решить поставленную задачу, введём функцию $y = x^a$. Если $x = 1$, то $y = 1$. Нам же надо найти значение функции для $x = 1 + h$, т. е. мы должны дать аргументу приращение $\Delta x = h$ и найти Δy. Ввиду малости h можно вместо Δy вычислить dy. Так как $y = x^a$, то $dy = ax^{a-1}dx$. У нас $x = 1$ (заменяя Δy на $dy = y'dx$, мы должны вычислять y' при том значении аргумента, к которому придаётся приращение Δx). Значит, $dy = ah$, и мы приходим к приближённой формуле

$$(1+h)^a \approx 1 + ah.$$

Из этой формулы можно получить быстрый приближённый способ извлечения корней. Именно, пусть n — число натуральное и B мало сравнительно с A^n. Тогда

$$\sqrt[n]{A^n+B}=A\left(1+\frac{B}{A^n}\right)^{\frac{1}{n}}\approx A\left(1+\frac{B}{nA^n}\right)$$

и

$$\sqrt[n]{A^n+B}\approx A+\frac{B}{nA^{n-1}}.$$

Например,

$$\sqrt[3]{131}=\sqrt[3]{125+6}=5+\frac{6}{3\cdot 5^2}=5{,}08.$$

В действительности, $\sqrt[3]{125}=5{,}0788\ldots$

Пример 2. Вычислить $\operatorname{tg} 31°$. Здесь надо рассмотреть функцию $y=\operatorname{tg} x$. Она известна для $x=30°=\frac{\pi}{6}$. Нас интересует её приращение, вызываемое приращением $\Delta x=1°=\frac{\pi}{180}$. Вместо этого приращения вычислим дифференциал

$$dy=\frac{\Delta x}{\cos^2 x}=\frac{1}{\cos^2 30°}\cdot\frac{\pi}{180}=\frac{\pi}{135}=0{,}023\ldots$$

Так как

$$\operatorname{tg} 30°=\frac{\sqrt{3}}{3}=0{,}577\ldots,$$

то

$$\operatorname{tg} 31°=0{,}600\ldots$$

В действительности, $\operatorname{tg} 31°=0{,}60076\ldots$

Пример 3. Пусть $y=uv$, где u и v — функции аргумента x. Допустим, что x получил малое приращение dx. Тогда

$$dy=y'dx=(uv'+u'v)\,dx=u\,dv+v\,du,$$

откуда

$$\frac{dy}{y}=\frac{du}{u}+\frac{dv}{v}.$$

Если u и v суть величины, измеряемые каким-либо прибором, а du и dv — ошибки измерения, то полученное равенство означает, что *относительная ошибка произведения равна сумме относительных ошибок сомножителей.*

В заключение установим, что дифференциал dy функции y полностью определяется двумя свойствами: пропорциональностью Δx и стремлением к нулю отношения

$$\frac{\Delta y-dy}{\Delta x}$$

при стремлении к нулю Δx. В самом деле, пусть z есть величина, обладающая этими свойствами, т. е. $z = k\Delta x$ и

$$\lim_{\Delta x \to 0} \frac{\Delta y - z}{\Delta x} = 0. \tag{14}$$

Так как

$$\frac{\Delta y - z}{\Delta x} = \frac{\Delta y}{\Delta x} - k,$$

то из (14) следует, что

$$k = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$

т. е. $k = y'$ и $z = dy$.

Таким образом, всякое «округление» Δy, при котором Δy заменяется (с точностью до малых порядка высшего относительно Δx) величиной, пропорциональной Δx, означает замену Δy дифференциалом dy.

Например, в курсе физики объёмный коэффициент расширения β какого-либо материала подсчитывается так: пусть куб из этого материала имеет сторону, равную 1, и нагревается на 1°. Тогда сторона куба становится равной $1 + \alpha$, где α — линейный коэффициент расширения материала. Поэтому объём нагретого куба равен

$$(1 + \alpha)^3 = 1 + 3\alpha + 3\alpha^2 + \alpha^3,$$

а увеличение объёма равно

$$3\alpha + 3\alpha^2 + \alpha^3.$$

Отбрасывая члены $3\alpha^2$ и α^3, что возможно в силу малости α, мы приходим к формуле

$$\beta = 3\alpha.$$

Читателю ясно, что суть этого рассуждения и состоит в замене приращения объёма дифференциалом этого объёма.

С идеей считать малые приращения функции пропорциональными приращениям аргумента читатель должен быть хорошо знаком также из практики работы с таблицами логарифмов. Таким образом, не говоря этого явно, в курсе средней школы неоднократно пользуются дифференциалами.

8. Производные и дифференциалы высшего порядка. Пусть функция $y = f(x)$ задана в каком-либо промежутке $[a, b]$ и в каждой точке [1]) этого промежутка имеет производную $y' = f'(x)$. Тогда эта производная сама является некоторой функцией от x (где x — точка дифференцирования). Поэтому можно говорить о производной

---

[1]) Говоря о производной в точках a и b, мы имеем в виду односторонние производные.

от $y'=f'(x)$. Эта последняя производная называется *второй производной* (или производной второго порядка) от функции $y=f(x)$ и обозначается через y'' или $f''(x)$.

Более общим образом производной порядка $n+1$ от функции $y=f(x)$ называется производная от производной порядка n. Для производной 3-го порядка применяется обозначение y''' или $f'''(x)$, дальнейшие же производные обозначаются через $y^{(IV)}$, $y^{(V)}$, ..., $y^{(n)}$, ..., или через $f^{(4)}(x)$, $f^{(5)}(x)$, ..., $f^{(n)}(x)$, ... Разумеется, не всякая функция обязательно имеет производную высшего порядка; кроме того, в одних точках такая производная может существовать, а в других — нет. Важно подчеркнуть, однако, что, предполагая существование производной $f^{(n+1)}(x)$ в какой-либо точке x_0, мы тем самым уже предполагаем существование предыдущей производной $f^{(n)}(x)$ не только в точке x_0, но и в некотором открытом [1]) промежутке, содержащем эту точку.

Хотя по самому определению производной порядка $n+1$ для её нахождения надо знать производную порядка n, но в некоторых простых случаях можно написать выражение производной высшего порядка, минуя все предыдущие. Например, совершенно ясно, что для $y=e^x$ будет $y^{(n)}=e^x$. Точно так же, замечая, что для $y=\sin x$ будет $y'=\cos x=\sin\left(x+\frac{\pi}{2}\right)$, мы видим, что здесь

$$y^{(n)}=\sin\left(x+n\frac{\pi}{2}\right).$$

Важным фактом является также то, что для многочлена степени m

$$y=a_0x^m+a_1x^{m-1}+\ldots+a_{m-1}x+a_m$$

будет $y^{(m)}=a_0m!$, а все производные порядка более высокого, чем m тождественно равны нулю.

Дадим здесь применение производных к выводу бинома Ньютона. Заметим, что если n — натуральное число, то $(a+x)^n$ есть многочлен степени n; это вытекает непосредственно из того, что степени многочленов (в данном случае двучленов) при перемножении складываются. Поэтому

$$(a+x)^n=A_0+A_1x+A_2x^2+\ldots+A_nx^n. \qquad (15)$$

Задача заключается в определении коэффициентов A_0, A_1, A_2, ..., A_n. Полагая $x=0$, получаем:

$$a^n=A_0.$$

---

[1]) Впрочем, этот промежуток может быть замкнутым и иметь точку x_0 одним из своих концов, но тогда $f^{(n+1)}(x_0)$ будет обозначать одностороннюю производную.

Если же тождество (15) продифференцировать, то получится новое тождество:

$$n(a+x)^{n-1} = A_1 + 2A_2x + \ldots + nA_nx^{n-1}. \qquad (16)$$

Полагая здесь $x = 0$, найдём:

$$na^{n-1} = A_1.$$

Дифференцируя тождество (16) ещё раз, получим:

$$n(n-1)(a+x)^{n-2} = 2 \cdot 1A_2 + 3 \cdot 2A_3x + \ldots + n(n-1)A_nx^{n-2}, \qquad (17)$$

откуда при $x = 0$:

$$n(n-1)a^{n-2} = 2 \cdot 1A_2.$$

Заметим, что равенство (16) получилось путём однократного, а равенство (17) путём двукратного дифференцирования равенства (15). Так как при каждом дифференцировании степени появляется новый коэффициент, равный исходному показателю степени, а сам показатель степени понижается на единицу, то после m-кратного дифференцирования равенства (15) получим ($m \leqslant n$):

$$\underbrace{n(n-1)\ldots(n-m+1)}_{m \text{ множителей}}(a+x)^{n-m} = m(m-1)\ldots 1 \cdot A_m +$$

$$+ (m+1)m\ldots 2 \cdot A_{m+1}x + \ldots + n(n-1)\ldots(n-m+1)A_nx^{n-m}.$$

Здесь первое слагаемое правой части является производной порядка m от A_mx^m, второе слагаемое — производной порядка m от $A_{m+1}x^{m+1}$ и т. д. Полагая в найденном тождестве $x = 0$, будем иметь:

$$n(n-1)\ldots(n-m+1)a^{n-m} = m(m-1)\ldots 1 \cdot A_m.$$

Так шаг за шагом, посредством всё новых и новых дифференцирований и затем подстановки частного значения $x = 0$, определяются все числа $A_0, A_1, A_2, \ldots, A_m, \ldots$ Мы получим, очевидно:

$$A_0 = a^n, \quad A_1 = na^{n-1}, \quad A_2 = \frac{n(n-1)}{2 \cdot 1}a^{n-2}, \ldots$$

$$\ldots, A_m = \frac{n(n-1)\ldots(n-m+1)}{m(m-1)\ldots 1}a^{n-m}, \ldots$$

$$\ldots, A_n = \frac{n(n-1)\ldots(n-n+1)}{n(n-1)\ldots 1}a^{n-n} = 1.$$

Поэтому тождество (15) приобретает вид

$$(a+x)^n = a^n + na^{n-1}x + \frac{n(n-1)}{1 \cdot 2}a^{n-2}x^2 +$$

$$+ \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3}x^3 + \ldots + x^n.$$

Это и есть бином Ньютона.

Таким образом, не прибегая к комбинаторике, бином Ньютона можно получить в качестве одного из применений производных.

Вообще дифференцирование является мощным средством для вывода тождеств. Так, например, из тождества

$$1+x+x^2+\ldots+x^n=\frac{1-x^{n+1}}{1-x}$$

посредством однократного дифференцирования получаем новое тождество:

$$1+2x+3x^2+\ldots+nx^{n-1}=\frac{1-(n+1)x^n+nx^{n+1}}{(1-x)^2}.$$

С понятием производной высшего порядка связано понятие дифференциала высшего порядка. Мы видели, что дифференциал $dy=y'dx$ зависит от двух величин — точки дифференцирования x и приращения $dx=\Delta x$ (в настоящий момент мы считаем x независимой переменной). Закрепив dx, мы превращаем dy в функцию одного аргумента x. Дифференциал этой новой функции и называется *вторым дифференциалом* (или дифференциалом второго порядка) функции y. Он сбозначается через d^2y [или, если $y=f(x)$, — через $d^2f(x)$]. Согласно определению для нахождения этого дифференциала надо найти производную функции dy и результат умножить на dx. Этот последний множитель dx, строго говоря, мог бы быть и отличен от того множителя dx, который участвовал в составлении дифференциала $dy=y'dx$. Однако (это новое соглашение!) эти множители выбираются равными. Поэтому

$$d^2y=d(dy)=d(y'dx)=(y'dx)'dx=y''(dx)^2.$$

Квадрат $(dx)^2$ принято обозначать коротко через dx^2 (хотя это обозначение иной раз может и привести к недоразумениям, ибо dx^2 можно принять за $d(x^2)$, т. е. за дифференциал функции x^2). Поэтому окончательно:

$$d^2y=y''dx^2.$$

Дифференциал этой функции (где снова dx закрепляется и вновь вводимый множитель dx принимается равным закреплённому) называется третьим дифференциалом функции y и обозначается через d^3y. Ясно, что

$$d^3y=y'''dx^3.$$

Аналогично этому, дифференциалом порядка $n+1$ называется дифференциал от дифференциала порядка n, причём попрежнему dx закреплён и взят одним и тем же для всех дифференциалов от 1-го до $(n+1)$-го порядков.

Легко понять, что

$$d^n y = y^{(n)} dx^n, \tag{18}$$

откуда вытекает возможность обозначать n-ю производную $y^{(n)}$ через

$$\frac{d^n y}{dx^n}.$$

Когда мы говорили о дифференциале dy, то подчеркнули инвариантный характер формулы

$$dy = y' dx,$$

состоящий в том, что она верна и тогда, когда x есть функция другого аргумента t. Нетрудно видеть, что для дифференциалов высшего порядка такой инвариантности быть не может, т. е. что формула (18) не будет верна, когда x является функцией от t. В самом деле, говоря уже о втором дифференциале d^2y, мы определяли его, как дифференциал выражения $dy = y'dx$, которое рассматривалось как функция от x при закреплённом dx. Но если x есть функция от t, то $dx = x'_t dt$ и закрепление dx (при изменении t) невозможно.

Высказанные соображения легко подтвердить и прямым вычислением. Действительно, если $y = f(x)$ и $x = \varphi(t)$, то y есть сложная функция от t. Значит,

$$d^2y = y''_t dt^2.$$

По правилу дифференцирования сложной функции

$$y''_t = (y'_x)'_t = (y'_x x'_t)'_t = (y'_x)'_t x'_t + y'_x x''_t.$$

Далее, $(y'_x)'_t = (y'_x)'_x \cdot x'_t = y''_x x'_t$. Поэтому

$$d^2y = y''_x x'^2_t dt^2 + y'_x x''_t dt^2.$$

Замечая, что

$$x'_t dt = dx, \quad x''_t dt^2 = d^2x,$$

приходим к формуле

$$d^2y = y''_x dx^2 + y'_x d^2x,$$

отличной от формулы (18).

9. Частные производные и полный дифференциал. Основное содержание настоящей статьи посвящено функциям одной переменной. Однако целесообразно сообщить некоторые сведения о дифференцировании функций нескольких переменных. Этот вопрос почти непосредственно приводится к дифференцированию функций одной переменной, ибо, если у функции нескольких аргументов закрепить все эти аргументы, кроме одного, то она превратится в функцию этого одного, незакреплённого, аргумента.

Пусть, например, $z = f(x, y)$ есть функция аргументов x и y. Закрепим точку (x_0, y_0) и найдём $f(x_0, y_0)$. Затем придадим аргументу x приращение

$\Delta x \neq 0$, *оставляя другой аргумент* y *закреплённым,* $y = y_0$. Новое значение функции будет равно $f(x_0 + \Delta x, y_0)$. Предел

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x} \tag{19}$$

называется *частной производной функции* $f(x, y)$ *по переменной* x *в точке* (x_0, y_0). Ясно, что это — обычная производная функции (одного аргумента!) $f(x, y_0)$. Обозначается предел (19) одним из символов

$$f'_x(x_0, y_0), \quad \frac{\partial f}{\partial x}, \quad z'_x, \quad \frac{\partial z}{\partial x}.$$

Аналогично определяется частная производная $\dfrac{\partial f}{\partial y}$.

Полезно подчеркнуть, что символ $\dfrac{\partial z}{\partial x}$ не есть частное, как это было в случае производной функции одной переменной. Сами по себе выражения dz и dx лишены смысла. Покажем на примере, что пренебрежение этим указанием может привести к недоразумениям.

Пусть переменные x, y и z связаны соотношением

$$xyz = 1. \tag{20}$$

Тогда $z = \dfrac{1}{xy}$ и $\dfrac{\partial z}{\partial x} = -\dfrac{1}{x^2 y}$. Аналогично $\dfrac{\partial x}{\partial y} = -\dfrac{1}{y^2 z}$ и $\dfrac{\partial y}{\partial z} = -\dfrac{1}{z^2 x}$. Отсюда $\dfrac{\partial x}{\partial y} \cdot \dfrac{\partial y}{\partial z} \cdot \dfrac{\partial z}{\partial x} = \dfrac{-1}{x^3 y^3 z^3}$ и в силу (20)

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = -1,$$

в то время, как, если бы написанные выражения были настоящими дробями, их произведение равнялось бы $+1$.

Читатель легко сообразит, каков геометрический смысл частных производных функции $z = f(x, y)$. Это — угловые коэффициенты касательных к сечениям поверхности $z = f(x, y)$ плоскостями $y = y_0$ и $x = x_0$.

Интересно, что, в отличие от функций одной переменной, существование обеих частных производных $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$ в какой-нибудь точке (x_0, y_0) ещё не обеспечивает непрерывности функции $f(x, y)$ в этой точке. Пусть, например, $f(x, y)$ задана так:

$$f(x, y) = \begin{cases} 0 & \text{при } xy \neq 0, \\ 1 & \text{при } xy = 0. \end{cases}$$

Эта функция разрывна в точке (0, 0), но обе частные производные в этой точке существуют.

Если частная производная $\dfrac{\partial z}{\partial x}$ существует всюду, то она сама является функцией от x и y и можно искать её частные производные. Это будут частные производные второго порядка, которые обозначаются так:

$$\frac{\partial^2 z}{\partial x^2}, \quad \frac{\partial^2 z}{\partial x \, \partial y}.$$

Следует различать символы $\dfrac{\partial^2 z}{\partial x \, \partial y}$ и $\dfrac{\partial^2 z}{\partial y \, \partial x}$. Однако в том случае, когда эти производные непрерывны, они равны

$$\frac{\partial^2 z}{\partial x \, \partial y} = \frac{\partial^2 z}{\partial y \, \partial x}.$$

К этим кратким замечаниям следует добавить ещё несколько соображений. Именно, как мы видели выше, приращение Δy дифференцируемой функции $y=f(x)$, вызванное приращением Δx аргумента x, можно представить в форме

$$\Delta y = f'(x)\,\Delta x + \rho, \tag{21}$$

где ρ стремится к нулю быстрее, чем Δx. Нечто подобное имеет место и для функций нескольких переменных. Пусть, например, $z=f(x, y)$ есть функция аргументов x и y, имеющая частные производные $f'_x(x, y)$ и $f'_y(x, y)$ во всех точках некоторого квадрата $a<x<b$, $c<y<d$, и эти частные производные сами являются непрерывными функциями точки дифференцирования (x, y). В этом случае приращение Δz функции z, вызванное приращениями Δx и Δy аргументов x и y, можно записать в форме

$$\Delta z = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y + \rho, \tag{22}$$

где x и y суть исходные значения аргументов, а ρ стремится к нулю быстрее, чем $|\Delta x|+|\Delta y|$. Доказательство этого утверждения будет приведено в п° 11. Выражение

$$dz = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y$$

называется *полным дифференциалом* функции $f(x, y)$. Оно является линейной функцией Δx и Δy и воспроизводит соответствующее им приращение Δz с точностью до бесконечно малой ρ порядка высшего, чем $|\Delta x|+|\Delta y|$. Таким образом, здесь имеется полная аналогия со случаем функции одного аргумента. Единственное отличие, которое здесь имеет место, состоит в том, что для справедливости формулы (21) достаточно было, чтобы производная $f'(x)$ существовала только при том значении аргумента x, к которому прибавляется приращение Δx, в случае же двух переменных существования частных производных $f'_x(x, y)$ и $f'_y(x, y)$ в одной лишь точке (x, y) недостаточно [1]) для справедливости формулы (22), а приходится требовать их существования и непрерывности в целом квадрате, содержащем эту точку внутри себя.

§ 2. Важнейшие теоремы о производных

10. Теоремы Ферма и Ролля. В п° 2 мы уже старались выяснить большое принципиальное значение понятия производной. В дальнейшем мы увидим, что с помощью этого понятия удаётся разрешать разнообразные и важные задачи конкретного характера. Однако для этой цели придётся изучить понятие производной более обстоятельно. Выяснению важнейших свойств производной и будет посвящён настоящий параграф.

Начнём с важной теоремы, обычно связываемой с именем П. Ферма [2]).

---

[1]) Как мы видели, функция $f(x, y)$ может в точке (x, y) иметь обе производные $f'_x(x, y)$ и $f'_y(x, y)$ и в то же время быть разрывной. Ясно, что в этом случае формула (22) не справедлива.

[2]) В действительности, Ферма не вводил понятия производной. Настоящая теорема представляет собой уточнение не вполне отчётливых соображений Ферма.

Теорема Ферма. *Пусть функция $f(x)$ задана в замкнутом промежутке $[a, b]$ и в некоторой внутренней* [1]) *точке x_0 принимает своё наибольшее (или наименьшее) значение. Если в этой точке существует производная $f'(x_0)$, то она обязательно равна нулю*

$$f'(x_0) = 0.$$

Доказательство. Пусть для определённости $f(x_0)$ есть наибольшее из значений функции на $[a, b]$. Так как точка x_0 есть *внутренняя* точка промежутка $[a, b]$, то производная $f'(x_0)$ одновременно является и правосторонней и левосторонней. Вычисляя её как правостороннюю, будем иметь

$$f'(x_0) = \lim \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x},$$

где $\Delta x > 0$ и $\Delta x \to 0$. Но из того, что $f(x_0)$ есть наибольшее значение $f(x)$, вытекает, что

$$f(x_0 + \Delta x) - f(x_0) \leqslant 0.$$

Стало быть, при $\Delta x > 0$ будет

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \leqslant 0.$$

Но тогда предел этого выражения не может быть положительным, и потому

$$f'(x_0) \leqslant 0.$$

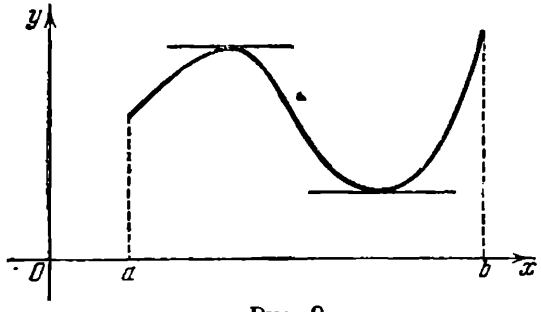


Рис. 9.

Аналогично, вычисляя $f'(x_0)$ как левостороннюю производную, мы установим, что

$$f'(x_0) \geqslant 0.$$

Теорема доказана.

Геометрически эта теорема означает, что, если в самой высокой и в самой низкой точках графика функции существуют касательные и если абсциссы этих точек лежат внутри промежутка задания функции, то упомянутые касательные параллельны оси Ox (рис. 9).

Читатель должен обратить внимание на то, что условие $a < x_0 < b$ весьма существенно, ибо важно, чтобы $f'(x_0)$ можно было рассмат-

[1]) То-есть такой, что $a < x_0 < b$.

ривать и как правостороннюю и как левостороннюю производные. Если же точка x_0, в которой $f(x)$ принимает наибольшее значение, совпадает с a или b, то при наличи $f'(x_0)$ нельзя гарантировать, что $f'(x_0)=0$. Иллюстрацией этого является рис. 10, на котором касательные в точках A и B не параллельны оси Ox.

При помощи теоремы Ферма доказывается следующее предложение, являющееся одной из основных теорем математического анализа.

Теорема Ролля. *Пусть функция $f(x)$ задана в замкнутом промежутке $[a, b]$ и во всех его точках имеет производную $f'(x)$. Если на концах промежутка $[a, b]$ значения функции равны между собой, т. е.*

$$f(a)=f(b),$$

то внутри $[a, b]$ обязательно существует хотя бы одна такая точка c $(a<c<b)$, что значение производной $f'(x)$ в этой точке равно нулю

$$f'(c)=0.$$

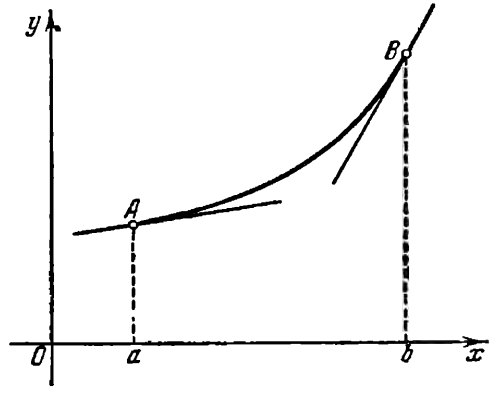


Рис. 10.

Если, в частности, предположить, что значения функции в точках a и b равны нулю, то теорему Ролля можно будет коротко формулировать так: *между двумя нулями функции содержится нуль её производной.*

Переходя к доказательству теоремы, заметим прежде всего, что функция $f(x)$ необходимо будет непрерывной, поскольку она имеет производную $f'(x)$. Но тогда среди её значений обязательно имеются и наибольшее и наименьшее (см. стр. 218).

Обозначим их через M и m. Если $M=m$, то это означает, что функция $f(x)$ постоянна. Но тогда во всех точках $[a, b]$ будет $f'(x)=0$ и любая из точек открытого промежутка (a, b) может быть принята за требуемую точку c. Если же $M>m$, то эти значения не могут оба достигаться на концах $[a, b]$, ибо по условию $f(a)=f(b)$. Стало быть, хоть одно из них принимается функцией $f(x)$ в точке x_0, лежащей внутри $[a, b]$, т. е. такой, что $a<x_0<b$. Но тогда по теореме Ферма будет $f'(x_0)=0$, т. е. x_0 и является требуемой точкой c.

Доказанная теорема является типичной «теоремой существования». В ней утверждается существование точки c с определёнными свойствами, но нет речи о каких бы то ни было способах нахождения этой точки.

11. Формулы Лагранжа и Коши. Правило Лопиталя. Следующие две теоремы представляют собой последовательные обобщения теоремы Ролля.

Теорема Лагранжа. *Пусть $f(x)$ — функция, заданная в замкнутом промежутке $[a, b]$ и имеющая во всех точках этого промежутка производную $f'(x)$. Тогда внутри $[a, b]$ обязательно существует хотя бы одна такая точка c, $a<c<b$, что*

$$\frac{f(b)-f(a)}{b-a}=f'(c). \tag{1}$$

Равенство (1) называется *формулой Лагранжа*, или *формулой конечных приращений*. Ясно, что эта формула превращается в равенство $f'(c)=0$ в случае, когда $f(a)=f(b)$, так что теорема Лагранжа действительно является обобщением теоремы Ролля. В свою очередь обобщением теоремы Лагранжа является следующая

Теорема Коши. *Пусть в промежутке $[a, b]$ заданы две функции $f(x)$ и $g(x)$, имеющие всюду на $[a, b]$ производные $f'(x)$ и $g'(x)$, причём $g'(x)$ не обращается в нуль ни в одной точке открытого промежутка (a, b). В таком случае между a и b обязательно существует такая точка c, $a<c<b$,* что

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}. \tag{2}$$

Формула (2) называется *формулой Коши*, или *обобщённой формулой конечных приращений*. Теорема Лагранжа получается из теоремы Коши в том случае, когда $g(x)=x$.

Очевидно, достаточно доказать формулу Коши. С этой целью введём вспомогательную функцию

$$\varphi(x)=[f(b)-f(a)][g(x)-g(a)]-[f(x)-f(a)][g(b)-g(a)].$$

Легко видеть, что всюду на $[a, b]$ у $\varphi(x)$ существует производная $\varphi'(x)$, причём

$$\varphi'(x)=[f(b)-f(a)]g'(x)-f'(x)[g(b)-g(a)].$$

Кроме того, непосредственно из выражения $\varphi(x)$ ясно, что $\varphi(a)=0$ и $\varphi(b)=0$. Таким образом, $\varphi(x)$ удовлетворяет всем условиям теоремы Ролля. Значит, в (a, b) существует такая точка c, что $\varphi'(c)=0$, или, что то же самое,

$$[f(b)-f(a)]g'(c)=f'(c)[g(b)-g(a)]. \tag{3}$$

Чтобы перейти отсюда к формуле (2), достаточно разделить (3) на $g'(c)[g(b)-g(a)]$. Последняя операция законна, если это произведение отлично от нуля. Но это и в самом деле так, ибо $g'(x)$ не обращается в нуль нигде в (a, b) и, в частности, $g'(c)\neq 0$. С другой стороны, если бы было $g(b)-g(a)=0$, то по теореме Ролля между a и b существовал бы корень производной $g'(x)$,

a так как такового корня нет, то и второй множитель $g(b)-g(a)$ отличен от нуля. Теорема доказана.

Если заметить, что отношение

$$\frac{f(b)-f(a)}{b-a}$$

представляет собой угловой коэффициент хорды графика функции $f(x)$, соединяющей точки $A[a, f(a)]$ и $B[b, f(b)]$ этого графика, то легко увидеть, что геометрический смысл теоремы Лагранжа таков [1]: *если кривая $y=f(x)$ всюду имеет касательную, то на любой дуге $\overset{\smile}{AB}$ этой кривой существует точка C, в которой касательная параллельна хорде AB* (рис. 11).

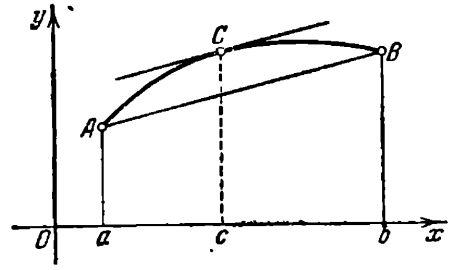


Рис. 11.

Формула Лагранжа весьма полезна для вывода неравенств и для приближённых вычислений. Приведём два примера.

Пример 1. Применим формулу (1) к функции x^α; получим:

$$\frac{b^\alpha-a^\alpha}{b-a}=\alpha c^{\alpha-1}, \quad 0<a<c<b.$$

Если $\alpha>1$, то $a^{\alpha-1}<c^{\alpha-1}<b^{\alpha-1}$, поэтому

$$\alpha a^{\alpha-1}<\frac{b^\alpha-a^\alpha}{b-a}<\alpha b^{\alpha-1},$$

откуда

$$\alpha a^{\alpha-1}(b-a)<b^\alpha-a^\alpha<\alpha b^{\alpha-1}(b-a).$$

Если $0<\alpha<1$, то $a^{\alpha-1}>c^{\alpha-1}>b^{\alpha-1}$, поэтому

$$\alpha b^{\alpha-1}<\frac{b^\alpha-a^\alpha}{b-a}<\alpha a^{\alpha-1},$$

откуда

$$\alpha b^{\alpha-1}(b-a)<b^\alpha-a^\alpha<\alpha a^{\alpha-1}(b-a).$$

Пример 2. Применим формулу (1) к функции $\ln x$, полагая $b=1+x$ $(x>0)$, $a=1$; получим:

$$\frac{\ln(1+x)-\ln 1}{(1+x)-1}=\frac{1}{c}, \quad \text{т. е.} \quad \ln(1+x)=\frac{x}{c},$$

где c лежит между 1 и $1+x$. Поэтому

$$\frac{x}{1+x}<\ln(1+x)<x.$$

[1]) В этом же состоит и геометрический смысл теоремы Коши, но, чтобы увидеть это, надо ввести кривую, заданную параметрическими уравнениями $x=f(t)$, $y=g(t)$.

Замечая, что $\lg(1+x)=0{,}43429...\ln(1+x)$ $(0{,}43429...=\lg e)$, умножим все члены найденных неравенств на $0{,}43429...$; получим:

$$\frac{0{,}43429...\,x}{1+x}<\lg(1+x)<0{,}43429...\,x.$$

Правая и левая части неравенств дают приближённые значения $\lg(1+x)$ по избытку и по недостатку; разность между ними есть $\frac{0{,}43429...\,x^2}{1+x}<0{,}5x^2$. Поэтому при $0<x<0{,}1$ получаем приближённое равенство с ошибкой, меньшей 0,005:

$$\lg(1+x)\approx 0{,}43x,$$

а при $0<x<0{,}01$ — приближённое равенство с ошибкой, меньшей 0,00005:

$$\lg(1+x)\approx 0{,}434x.$$

Ниже мы приведём несколько важных теоретических примеров приложений теорем Лагранжа и Коши. Здесь же рассмотрим ещё три вопроса сравнительно частного характера.

I. Пусть функция $y=f(x)$ всюду на $[a, b]$ имеет вторую производную $f''(x)$. Рассмотрим какое-нибудь значение $x\in[a, b]$, дадим ему приращение Δx и выясним, какую ошибку мы делаем, заменяя соответствующее приращение функции Δy дифференциалом $dy=y'\Delta x$.

По теореме Лагранжа

$$\Delta y=f(x+\Delta x)-f(x)=f'(x_1)\,\Delta x,$$

где $x<x_1<x+\Delta x$ (для определённости мы считаем $\Delta x>0$).

С другой стороны, $dy=f'(x)\,\Delta x$. Значит,

$$\Delta y-dy=[f'(x_1)-f'(x)]\,\Delta x.$$

Вторичное применение теоремы Лагранжа даёт

$$f'(x_1)-f'(x)=f''(x_2)\,(x_1-x)\qquad(x<x_2<x_1).$$

Допуская, что всюду в $[a, b]$ вторая производная $f''(x)$ по абсолютной величине не превосходит числа M, и замечая, что $|x_1-x|<\Delta x$, получаем оценку

$$|\Delta y-dy|\leqslant M(\Delta x)^2,$$

о которой мы уже упоминали в п° 7.

II. Допустим, что непрерывные функции $f(x)$ и $g(x)$ в одной и той же точке $x=a$ обращаются в нуль, $f(a)=g(a)=0$, но в точках, близких к a, функция $g(x)$ отлична от нуля. Отношение

$$\frac{f(x)}{g(x)}$$

представляет собой отношение двух величин, каждая из которых при приближении x к a стремится к нулю. Подобные выражения носят название «неопределённостей вида $\frac{0}{0}$», а отыскание предела такого выражения (или доказательство отсутствия этого предела) называется «раскрытием неопределённости». Теорема Коши даёт возможность решать задачу раскрытия неопределённости вида $\frac{0}{0}$ при помощи некоторого общего правила, обычно называемого *правилом Лопиталя*. В основе этого правила лежит

Теорема. *Пусть функции $f(x)$ и $g(x)$ удовлетворяют всем условиям теоремы Коши, и, кроме того, $f(a)=g(a)=0$. Если существует (конечный или бесконечный) предел отношения производных*

$$\lim_{x\to a}\frac{f'(x)}{g'(x)}=A,$$

то к тому же пределу стремится и отношение функций, т. е.

$$\lim_{x\to a}\frac{f(x)}{g(x)}=A. \tag{4}$$

В самом деле, закрепляя точку x, для которой $a<x\leqslant b$, и применяя теорему Коши к промежутку $[a, x]$, находим:

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(x_1)}{g'(x_1)},$$

где $a<x_1<x$. Если $x\to a$, то и подавно $x_1\to a$ и потому

$$\frac{f'(x_1)}{g'(x_1)}\to A,$$

откуда и следует (4).

Приведём два примера применения правила Лопиталя:

$$1)\ \lim_{x\to 3}\frac{x^3-27}{x^2-4x+3}=\lim_{x\to 3}\frac{3x^2}{2x-4}=\frac{27}{2},$$

$$2)\ \lim_{x\to 0}\frac{x-\sin x}{x^3}=\lim_{x\to 0}\frac{1-\cos x}{3x^2}.$$

Выражение $\frac{1-\cos x}{3x^2}$ само есть неопределённость вида $\frac{0}{0}$. Вторично применяя правило Лопиталя, находим:

$$\lim_{x\to 0}\frac{x-\sin x}{x^3}=\lim_{x\to 0}\frac{\sin x}{6x}=\frac{1}{6}.$$

III. Вернёмся к вопросу о полном дифференциале, затронутому в конце п° 9, и докажем формулу (22), приведённую там без доказательства.

Пусть функция $z=f(x, y)$ задана в квадрате Q, определённом неравенствами $a<x<b$, $c<y<d$, имеет во всех точках этого квадрата частные

производные $f'_x(x, y)$ и $f'_y(x, y)$, причём эти производные сами суть непрерывные функции. Закрепим точку (x, y) из квадрата Q и пусть $(x+\Delta x, y+\Delta y)$ есть другая точка этого квадрата. Рассмотрим разность

$$\Delta z = f(x+\Delta x, y+\Delta y) - f(x, y).$$

Эту разность можно переписать так:

$$\Delta z = [f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)] + [f(x, y+\Delta y) - f(x, y)].$$

Разность $f(x, y+\Delta y) - f(x, y)$, стоящая во вторых квадратных скобках, есть приращение функции $f(x, y)$, рассматриваемой как функция одного аргумента y (x закреплено). Значит, мы можем применить к этой разности формулу Лагранжа

$$f(x, y+\Delta y) - f(x, y) = f'_y(x, \eta)\,\Delta y,$$

причём η есть величина, содержащаяся между y и $y+\Delta y$. Аналогично

$$f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) = f'_x(\xi, y+\Delta y)\,\Delta x,$$

причём ξ лежит между x и $x+\Delta x$.

Заметим теперь, что в силу предположенной непрерывности обеих частных производных f'_x и f'_y каждая из разностей

$$\alpha = f'_x(\xi, y+\Delta y) - f'_x(x, y), \quad \beta = f'_y(x, \eta) - f'_y(x, y)$$

стремится к нулю при Δx и Δy, стремящихся к нулю [1]). Поэтому, положив

$$\alpha\Delta x + \beta\Delta y = \rho,$$

мы получим:

$$\Delta z = f'_x(x, y)\,\Delta x + f'_y(x, y)\,\Delta y + \rho,$$

причём

$$\frac{|\rho|}{|\Delta x| + |\Delta y|} \leqslant |\alpha| + |\beta|.$$

Ясно, что при Δx и Δy, стремящихся к нулю, окажется

$$\lim \frac{\rho}{|\Delta x| + |\Delta y|} = 0.$$

Этим и оправданы все утверждения [2]), сделанные нами в п° 9.

12. Формула Тейлора. Общее определение функции гласит: «величина y является функцией аргумента x, если каждому значению x отвечает определённое значение y». В этом определении нет речи о том, при помощи каких средств можно найти значение y, соответствующее заданному значению x. Во многих случаях функция задаётся той или иной вычислительной формулой, например формулой

$$y = x^2.$$

---

[1]) Читатель обратит внимание на то, что при $\Delta x \to 0$, $\Delta y \to 0$ будет $\xi \to x$, $\eta \to y$.

[2]) Из доказательства видно, что хотя существование пронзводных f'_x и f'_y приходится предполагать имеющим место во всех точках Q, но непрерывность их нужна только в исходной точке (x, y).

В этих случаях сама формула и доставляет нам способ находить y по x. Но часто дело обстоит не так. Например, задавая функцию равенством

$$y = \sin x,$$

мы не получаем в руки никакого вычислительного (или, как чаще говорят, аналитического) способа находить y. Возникает важная и общая задача создания аналитических средств, позволяющих находить значения функции по заданным значениям аргумента. В широком классе случаев указанная задача решается при помощи замечательной формулы, называемой *формулой Тейлора*.

Займёмся сначала установлением этой формулы для того случая, когда рассматриваемая функция является многочленом

$$f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n. \tag{5}$$

В основе вывода формулы для этого случая лежит то простое замечание, что при любом действительном числе a функцию $f(x)$ можно записать в форме

$$f(x) = A_0 + A_1 (x-a) + A_2 (x-a)^2 + \ldots + A_n (x-a)^n, \tag{6}$$

где $A_0, A_1, \ldots, A_n$ не зависят от x. Иначе говоря, мы утверждаем, что многочлен (5) можно расположить по степеням разности $x-a$. Чтобы установить этот факт, достаточно доказать его для функции x^k, ибо $f(x)$ есть сумма таких функций, умноженных на постоянные коэффициенты C_k. Для функции же x^k наше утверждение вытекает из того, что

$$x^k = [a + (x-a)]^k = \sum_{i=0}^{k} C_k^i a^i (x-a)^{k-i}.$$

Постараемся теперь фактически найти коэффициенты $A_0, A_1, \ldots, A_n$, входящие в формулу (6). Коэффициент A_0 находится сразу, если положить в этой формуле $x = a$, что даёт

$$A_0 = f(a).$$

Чтобы найти следующий коэффициент A_1, возьмём от обеих частей равенства (6) производные, что даёт

$$f'(x) = A_1 + 2A_2 (x-a) + 3A_3 (x-a)^2 + \ldots + nA_n (x-a)^{n-1}. \tag{7}$$

Полагая здесь $x = a$, получаем:

$$A_1 = f'(a).$$

Дифференцируя равенство (7), находим:

$$f''(x) = 2A_2 + 3 \cdot 2A_3 (x-a) + \ldots + n(n-1) A_n (x-a)^{n-2}.$$

Полагая $x = a$, получим: $A_2 = \dfrac{f''(a)}{1 \cdot 2}$.

Аналогично мы установим и общую формулу

$$A_k = \frac{f^{(k)}(a)}{k!}.$$

Таким образом, равенство (6) принимает вид

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \quad (8)$$

Это и есть формула Тейлора для многочлена.

Допустим теперь, что $f(x)$ есть произвольная функция, имеющая однако, в точке a все производные до n-й включительно. Тогда мы сможем для неё составить многочлен

$$T(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n, \quad (9)$$

стоящий в правой части формулы (8). Однако мы не можем уже утверждать, что этот многочлен совпадает с функцией $f(x)$.

Тем не менее в очень большом числе случаев имеет место замечательный факт, что *разность между функцией $f(x)$ и многочленом $T(x)$ весьма мала.* Всякий раз, когда это так, мы получаем в многочлене $T(x)$ хорошее приближённое выражение для функции $f(x)$. Более того, оказывается, что в том случае, когда $f(x)$ имеет производные *всех* порядков и мы можем число n брать сколь угодно большим, то весьма часто с увеличением этого числа упомянутая разность $f(x) - T(x)$ *стремится к нулю,* так что мы получаем возможность вычислять $f(x)$ с любой степенью точности.

Переходя к точному изложению вопроса, предположим, что $f(x)$ задана в некотором промежутке $[A, B]$, содержащем точку a, и имеет во всех точках этого промежутка производные всех порядков до $(n+1)$-го включительно[1]). Введём многочлен $T(x)$, определив его формулой (9). Тогда

$$T'(x) = f'(a) + \frac{f''(a)}{1!}(x-a) + \ldots + \frac{f^{(n)}(a)}{(n-1)!}(x-a)^{n-1},$$

$$T''(x) = f''(a) + \frac{f''(a)}{1!}(x-a) + \ldots + \frac{f^{(n)}(a)}{(n-2)!}(x-a)^{n-2},$$

. . . . . .

$$T^{(n)}(x) = f^{(n)}(a),$$

$$T^{(n+1)}(x) = 0.$$

---

[1]) Если a совпадает с A или B, то под выражениями $f'(a)$, $f''(a)$, ..., $f^{(n)}(a)$, появляющимися ниже, разумеются односторонние производные.

Отсюда и из (9) следует, что

$$\left.\begin{array}{c} T(a)=f(a), \quad T'(a)=f'(a), \\ T''(a)=f''(a), \ \ldots, \quad T^{(n)}(a)=f^{(n)}(a), \quad T^{(n+1)}x=0. \end{array}\right\} \qquad (10)$$

Введём функцию $\varphi(x)$, положив

$$\varphi(x)=f(x)-T(x).$$

Из (10) следует, что

$$\varphi(a)=\varphi'(a)=\varphi''(a)=\ldots=\varphi^{(n)}(a)=0, \qquad \varphi^{(n+1)}(x)=f^{(n+1)}(x).$$

Введём, наконец, ещё одну вспомогательную функцию $\psi(x)$, положив

$$\psi(x)=\frac{(x-a)^{n+1}}{(n+1)!}.$$

Нетрудно видеть, что

$$\psi(a)=\psi'(a)=\psi''(a)=\ldots=\psi^{(n)}(a)=0, \quad \psi^{(n+1)}(x)=1.$$

Отметим также, что ни $\psi(x)$, ни одна из её производных (до $(n+1)$-й включительно) не обращаются в нуль в точках, отличных от a.

Заметив это, закрепим $x\in[A, B]$, считая $x\neq a$. Так как $\varphi(a)=\psi(a)=0$, то

$$\frac{\varphi(x)}{\psi(x)}=\frac{\varphi(x)-\varphi(a)}{\psi(x)-\psi(a)}.$$

К последнему выражению можно применить формулу Коши, что даёт

$$\frac{\varphi(x)}{\psi(x)}=\frac{\varphi'(x_1)}{\psi'(x_1)},$$

где x_1 лежит между точками a и x. Но $\varphi'(a)=\psi'(a)=0$ и потому

$$\frac{\varphi'(x_1)}{\psi'(x_1)}=\frac{\varphi'(x_1)-\varphi'(a)}{\psi'(x_1)-\psi'(a)}.$$

Вторичное применение формулы Коши даёт

$$\frac{\varphi(x)}{\psi(x)}=\frac{\varphi''(x_2)}{\psi''(x_2)},$$

где x_2 лежит между a и x_1, и тем более между a и x.

Продолжая это рассуждение, мы приходим к равенству

$$\frac{\varphi(x)}{\psi(x)}=\frac{\varphi^{(n+1)}(x_{n+1})}{\psi^{(n+1)}(x_{n+1})},$$

где x_{n+1} лежит между a и x.

Вспоминая, что

$$\varphi^{(n+1)}(x)=f^{(n+1)}(x),\ \psi^{(n+1)}(x)=1,$$

и обозначая x_{n+1} через $\bar{x}$, находим:

$$\frac{\varphi(x)}{\psi(x)}=f^{(n+1)}(\bar{x}),$$

или, что то же самое,

$$f(x)-T(x)=\frac{f^{(n+1)}(\bar{x})}{(n+1)!}(x-a)^{n+1}.$$

Подставляя сюда выражение (9) для $T(x)$, мы и приходим *к формуле Тейлора с остаточным членом*:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+$$
$$+\frac{f^{(n+1)}(\bar{x})}{(n+1)!}(x-a)^{n+1}, \quad (11)$$

причём (неизвестная нам) точка $\bar{x}$ лежит между a и x.

Таким образом, принимая многочлен $T(x)$ за функцию $f(x)$, мы совершим ошибку, равную

$$R_n(x)=\frac{f^{(n+1)}(\bar{x})}{(n+1)!}(x-a)^{n+1}.$$

Остановимся на том, как меняется эта ошибка с изменением n.

Теорема. *Если функция $f(x)$ в промежутке $[A, B]$ имеет производные всех порядков и если существует такое постоянное число K, что при всех $x\in[A, B]$ и при всех натуральных n будет*

$$|f^{(n)}(x)|\leqslant K,$$

то при любом x из $[A, B]$

$$\lim_{n\to\infty} R_n(x)=0.$$

В самом деле, в условиях теоремы

$$|R_n(x)|\leqslant K\,\frac{M^{n+1}}{(n+1)!},$$

где положено для краткости $|x-a|=M$. Возьмём столь большое натуральное m, чтобы оказалось

$$m+1>2M.$$

Тогда при $n>m$ будет

$$\frac{M^{n+1}}{(n+1)!}=\frac{M^m}{m!}\cdot\frac{M}{m+1}\cdot\frac{M}{m+2}\cdots\frac{M}{n+1}.$$

Каждая из $n-m+1$ дробей

$$\frac{M}{m+1},\ \frac{M}{m+2},\ \ldots,\ \frac{M}{n+1}$$

меньше $\frac{1}{2}$. Значит,

$$|R_n(x)| \leqslant K\frac{M^m}{m!}\cdot\frac{1}{2^{n-m+1}},$$

или, что то же самое,

$$|R_n(x)| \leqslant 2^{m-1}K\frac{M^m}{m!}\cdot\frac{1}{2^n}.$$

Обозначая постоянное (ведь x закреплён!) число

$$2^{m-1}K\frac{M^m}{m!}$$

через C, получаем неравенство

$$|R_n(x)| \leqslant \frac{C}{2^n},$$

из которого и следует теорема.

Таким образом, в условиях теоремы

$$f(x) = \lim_{n\to\infty}\left[f(a)+\frac{f'(a)}{1!}(x-a)+\ \ldots\ +\frac{f^{(n)}(a)}{n!}(x-a)^n\right],$$

чем и оправданы те общие соображения, которые были высказаны в начале п°.

Приведём несколько примеров.

Пример 1. Так как функция $f(x)=\sin x$ удовлетворяет условиям теоремы и для неё

$$f'(x)=\cos x,\quad f''(x)=-\sin x,\quad f'''(x)=-\cos x,\quad f^{(4)}(x)=\sin x,$$

$$f^{(5)}(x)=\cos x \text{ и вообще } f^{(n)}(x)=\sin\left(x+n\frac{\pi}{2}\right),$$

то, приняв $a=0$, получаем приближённое равенство

$$\sin x \approx x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ \ldots\ +(-1)^p\frac{x^{2p+1}}{(2p+1)!},$$

точность которого неограниченно улучшается с увеличением p.

Ошибка этого равенства не превосходит числа

$$\frac{|x|^{2p+2}}{(2p+2)!}.$$

Например, ошибка равенства

$$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!} \tag{12}$$

не больше, чем

$$\frac{|x|^8}{8!}=\frac{|x|^8}{40\,320}.$$

Значит, при $0<x\leqslant\frac{\pi}{4}$ вычисление функции $\sin x$ по формуле (12) приводит к ошибке, меньшей чем[1]) 0,000004.

Пример 2. Аналогичным образом

$$\cos x\approx 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots+(-1)^p\frac{x^{2p}}{(2p)!}.$$

Мы видим, что формула Тейлора даёт нам средство фактически вычислять тригонометрические функции. В главе, посвящённой теории рядов, мы ещё вернёмся к затронутым здесь соображениям.

Пример 3. Рассмотрим функцию x^c, где $c>1$; при $x>1$, $a=1$ и $n=1$ формула (11) даёт:

$$x^c=1+c(x-1)+\frac{c(c-1)}{1\cdot 2}\bar{x}^{c-2}(x-1)^2,\quad 1<\bar{x}<x$$

(мы пользуемся тем, что $(x^c)'=cx^{c-1}$, $(x^c)''=c(c-1)x^{c-2}$). Так как $\bar{x}^{c-2}=\frac{\bar{x}^c}{\bar{x}^2}>\frac{1}{x^2}$, то

$$x^c>1+c(x-1)+\frac{c(c-1)}{2}\left(\frac{x-1}{x}\right)^2\quad(x>1).$$

Отсюда, в частности, вытекает менее сильное, но зато более простое неравенство:

$$x^c-1>c(x-1)\quad(x>1).$$

Заметим, что оно выведено нами для любого $c>1$ (вообще говоря, не целого).

Пример 4. На стр. 198 было строго показано, что $0<e\leqslant 3$ и приводилось численное значение постоянной e. Покажем, как находится это значение.

Применим формулу Тейлора к функции $f(x)=e^x$, взяв $a=0$. Так как здесь $f^{(n)}(x)=e^x$, то $f^{(n)}(0)=1$ и

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+e^{\bar{x}}\frac{x^{n+1}}{(n+1)!},$$

где $\bar{x}$ лежит между 0 и x. В частности, при $x=1$ имеем:

$$e=1+\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{n!}+\frac{e^{\bar{x}}}{(n+1)!}\qquad(0<\bar{x}<1).\quad(13)$$

---

[1]) Этот подсчёт произведён так:

$\lg\pi=0{,}49715$, $\lg 4=0{,}60206$, $\lg\frac{\pi}{4}=\bar{1}\,,89509$, $\lg\left(\frac{\pi}{4}\right)^8=\bar{1}{,}16072$.

Значит $\left(\frac{\pi}{4}\right)^8<0{,}15$ и

$$\frac{\left(\frac{\pi}{4}\right)^8}{40320}<\frac{1}{268800}<0{,}000004.$$

Возьмём здесь $n=8$. Так как

$$\frac{e^{\bar{x}}}{9!}<\frac{e}{9!}\leqslant\frac{3}{9!}=\frac{3}{362880}<0{,}0000\,1,$$

то с ошибкой, меньшей чем 0,00001, будет

$$e=1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\frac{1}{6!}+\frac{1}{7!}+\frac{1}{8!}=2{,}718275.$$

Значит,

$$2{,}718275<e<2{,}718285.$$

В действительности,

$$e=2{,}718281828459\ldots$$

Формула (13) позволяет также очень просто установить, что e *есть число иррациональное* [1]). В самом деле, допустим, что e равно рациональной дроби $\frac{p}{q}$. Возьмём в формуле (13) число n большим чем q и большим чем 2. Умножая эту формулу на $n!$, получим:

$$\frac{pn!}{q}=\left[1+\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{n!}\right]\cdot n!+\frac{e^{\bar{x}}}{n+1}.$$

Отсюда следует, что дробь $\frac{e^{\bar{x}}}{n+1}$ есть число *целое*. С другой стороны, $0<\bar{x}<1$. Значит, $1<e^{\bar{x}}<e<3$, и потому эта дробь лежит между числами $\frac{1}{n+1}$ и $\frac{3}{n+1}$ и тем более между числами 0 и 1, откуда ясно, что она не может быть числом целым.

13. Исследования П. Л. Чебышева и С. Н. Бернштейна. Рассматривая вопрос о близости тейлорова многочлена

$$T(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n$$

к функции $f(x)$, мы предполагали x закреплённым и увеличивали n. Если, напротив, закрепить число n, то, вообще говоря, многочлен $T(x)$ уже не будет близок к $f(x)$ при всех x из $[A, B]$. Лишь при x, близких к a, можно говорить о близости $T(x)$ к $f(x)$, ибо обе эти функции, будучи непрерывными, совпадают при $x=a$. Таким образом, при закреплённом n многочлен $T(x)$ будет доставлять нам лишь л о к а л ь н о е приближение к $f(x)$. Улучшение же свойств этого приближения за счёт увеличения n может оказаться практически не очень удобным, ибо обращение с многочленами высоких степеней обычно связано с довольно громоздкими вычислениями.

Эти соображения привели великого русского математика П. Л. Чебышева к задаче построения таких многочленов, степень

---

[1]) С помощью более сильных средств можно доказать, что *число e трансцендентно*, т. е. что оно не является корнем никакого алгебраического уравнения с целыми коэффициентами.

которых заранее ограничена и которые всё же, в отличие от тейлорова многочлена $T(x)$, доставляют хорошее приближение функции $f(x)$ во всём промежутке $[A, B]$. Точно проблема Чебышева формулируется так: на промежутке $[A, B]$ задана непрерывная функция $f(x)$; требуется выбрать из всех многочленов $P(x)$ степени не выше n тот, для которого величина

$$\max_{A \leqslant x \leqslant B} |f(x) - P(x)| \tag{14}$$

имеет наименьшее значение. Оказывается, что такой многочлен всегда существует и единственен. Он называется *многочленом, наименее отклоняющимся от функции* $f(x)$. Само же наименьшее значение величины (14) называется наименьшим отклонением многочленов степени не выше n от функции $f(x)$ и обозначается обычно через $E_n(f)$. Чебышев обстоятельно изучил свойства многочленов, наименее отклоняющихся от данных функций, и дал целый ряд практических приложений своей теории — в теории машин и механизмов, в картографии и др. Замечательные исследования Чебышева послужили исходным пунктом для обширного ряда работ его учеников — К. А. Поссе, А. Н. Коркина, Е. И. Золотарёва, А. А. Маркова, В. А. Маркова и других.

Существенное развитие идеи П. Л. Чебышева получили в исследованиях С. Н. Бернштейна. Исходная мысль здесь заключается в том, что с увеличением n величина наименьшего отклонения $E_n(f)$ уменьшается, стремясь к нулю. Естественно, что быстрота убывания $E_n(f)$ связана со свойствами функции $f(x)$: чем проще строение этой функции, тем с большей точностью её можно заменить многочленом. Эти общие соображения привели С. Н. Бернштейна к созданию стройной классификации непрерывных функций на основании быстроты убывания их наименьших отклонений $E_n(f)$. В настоящее время вся совокупность связанных с этими идеями вопросов разработана весьма обстоятельно и составляет предмет важной современной области анализа — конструктивной теории функций, имеющей большое прикладное значение. В развитии этой дисциплины ведущая роль принадлежит советским исследователям. Помимо С. Н. Бернштейна, крупные заслуги в этой области имеют А. Н. Колмогоров, В. Л. Гончаров, М. Г. Крейн, Н. И. Ахиезер, Е. Я. Ремез, С. М. Лозинский, С. М. Никольский и другие.

§ 3. Применение дифференциального исчисления к исследованию функций

14. Признаки постоянства и монотонности функции. В настоящем п° мы установим, как по свойствам производной какой-нибудь функции судить о свойствах самой этой функции.

Теорема 1. *Пусть на замкнутом промежутке* $[a, b]$ *задана функция* $f(x)$. *Для того, чтобы эта функция была постоянной,*

необходимо и достаточно, чтобы во всех точках $[a, b]$ *существовала производная* $f'(x)$ *и чтобы она всюду на* $[a, b]$ *была равна нулю.*

Необходимость условий теоремы очевидна, ибо постоянная величина имеет производную и эта производная равна нулю. Чтобы доказать достаточность условий, предположим, что всюду на $[a, b]$ существует $f'(x)$ и $f'(x)=0$. Возьмём на $[a, b]$ произвольную точку x и применим теорему Лагранжа к промежутку $[a, x]$:

$$f(x)-f(a)=f'(\bar{x})\ (x-a) \qquad (a<\bar{x}<x).$$

Так как $f'(\bar{x})=0$, то $f(x)=f(a)$, откуда и следует, что $f(x)$ есть величина постоянная.

Напомним, что функция $f(x)$ называется *неубывающей* [*невозрастающей*], если из неравенства $x<y$ следует, что $f(x)\leqslant f(y)$ $[f(x)\geqslant f(y)]$. Если же из того, что $x<y$ следует $f(x)<f(y)$ $[f(x)>f(y)]$, то говорят, что $f(x)$ — функция *возрастающая* [*убывающая*]. Функции невозрастающие и неубывающие называются *монотонными.*

Если функция $f(x)$ возрастает, то $-f(x)$ убывает, и наоборот. Это простое замечание позволит нам при доказательствах нижеследующих теорем ограничиться рассмотрением только возрастающих функций.

Теорема 2. *Пусть на замкнутом промежутке* $[a, b]$ *задана функция* $f(x)$, *имеющая во всех точках* $[a, b]$ *производную* $f'(x)$. *Для того чтобы функция* $f(x)$ *была неубывающей* (*невозрастающей*), *необходимо и достаточно, чтобы всюду в открытом промежутке* (a, b) *было*

$$f'(x)\geqslant 0 \qquad [f'(x)\leqslant 0].$$

Допустим, что функция $f(x)$ не убывает. Закрепим $x\in(a, b)$ и выберем столь малое $\Delta x>0$, чтобы было $x+\Delta x\in[a, b]$. Тогда $f(x+\Delta x)\geqslant f(x)$, стало быть,

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}\geqslant 0.$$

Переходя в этом неравенстве к пределу при $\Delta x\to 0$, мы и получаем, что $f'(x)\geqslant 0$.

Допустим теперь, что при всех $x\in(a, b)$ будет $f'(x)\geqslant 0$. Возьмём на $[a, b]$ точки x и y, где $x<y$ и применим теорему Лагранжа к промежутку $[x, y]$:

$$f(y)-f(x)=f'(z)\ (y-x).$$

Здесь $x<z<y$ и потому $z\in(a, b)$. Значит, $f'(z)\geqslant 0$ и

$$f(y)\geqslant f(x),$$

так что $f(x)$ — функция неубывающая.

Теорема 3. *Пусть на замкнутом промежутке* $[a, b]$ *задана функция* $f(x)$, *имеющая во всех точках* $[a, b]$ *производную* $f'(x)$. *Для того чтобы* $f(x)$ *была возрастающей (убывающей) функцией, необходимо и достаточно выполнение двух условий*:

1) *Всюду в открытом промежутке* (a, b) *оказывается*

$$f'(x) \geqslant 0 \quad [f'(x) \leqslant 0].$$

2) *Не существует промежутка* $[p, q]$, *содержащегося в* $[a, b]$, *во всех точках которого было бы* $f'(x) = 0$.

Пусть $f(x)$ — функция возрастающая. Тогда по предыдущей теореме выполнено условие 1). Чтобы доказать необходимость условия 2), допустим, что оно нарушается. Тогда существует такой промежуток $[p, q]$, содержащийся в $[a, b]$, во всех точках которого $f'(x) = 0$. Согласно теореме 1 функция $f(x)$ будет постоянной на $[p, q]$ и не будет возрастающей. Таким образом установлена необходимость обоих условий 1) и 2).

Допустим теперь, что выполнены условия 1) и 2). По предыдущей теореме функция $f(x)$ оказывается неубывающей. Убедимся, что она возрастает. Для этого возьмём точки x и y из $[a, b]$, причём $x < y$. Для любого z из $[x, y]$ будет $f(x) \leqslant f(z) \leqslant f(y)$. Если бы оказалось, что $f(x) = f(y)$, то предыдущее неравенство привело бы нас к тому, что наша функция постоянна на

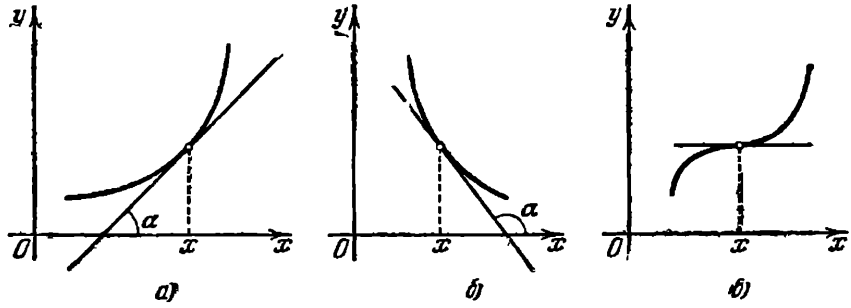


Рис. 12.

промежутке $[x, y]$. Но тогда её производная во всех точках этого промежутка обращалась бы в нуль, что противоречит условию 2). Стало быть, $f(x) \neq f(y)$ и $f(x) < f(y)$. Теорема доказана полностью.

Если вспомнить, что производная равна угловому коэффициенту касательной к графику функции, то доказанные результаты становятся геометрически почти очевидными. Действительно, график возрастающей функции при движении слева направо подымается (рис. 12, *а*), а график убывающей функции — опускается (рис. 12, *б*). Ясно, что в первом случае касательная к графику образует с осью

Ox острый угол, а во втором случае — тупой. Однако могут быть отдельные точки графика, в которых касательная параллельна оси Ox (рис. 12, *в*). Аналитически это означает, что даже и у возрастающей (убывающей) функции производная может в отдельных точках обращаться в нуль[1]).

Пример 1. Пусть

$$\varphi(x)=\frac{\sin x}{x}.$$

Так как эта формула теряет смысл при $x=0$, то мы будем рассматривать $\varphi(x)$ на промежутке $\left[\delta, \frac{\pi}{2}\right]$, где $0<\delta<\frac{\pi}{2}$. Она имеет здесь производную

$$\varphi'(x)=\frac{x\cos x-\sin x}{x^2}=\frac{x-\operatorname{tg} x}{x^2}\cos x.$$

Но при $0<x<\frac{\pi}{2}$ будет $x<\operatorname{tg} x$. Значит, в $\left(\delta, \frac{\pi}{2}\right)$ будет $\varphi'(x)<0$ и, стало быть, $\varphi(x)$ убывает на $\left[\delta, \frac{\pi}{2}\right]$. Отсюда $\varphi(x)\geqslant\varphi\left(\frac{\pi}{2}\right)=\frac{2}{\pi}$. Иначе говоря,

$$\sin x\geqslant\frac{2}{\pi}x.$$

Это неравенство доказано для $\delta\leqslant x\leqslant\frac{\pi}{2}$, но так как δ можно взять сколь угодно близким к нулю, то наше неравенство верно при $0<x\leqslant\frac{\pi}{2}$. Поскольку же оно превращается при $x=0$ в очевидное тождество $0=0$, то оно доказано для всех x из $\left[0, \frac{\pi}{2}\right]$.

Пример 2. Докажем, что при $p>1$ и любых положительных a и b справедливо неравенство

$$(a+b)^p\leqslant 2^{p-1}(a^p+b^p).$$

Предположим для определённости, что $a\leqslant b$, тогда $0<\frac{a}{b}\leqslant 1$. Разделим обе части предполагаемого неравенства на b^p и заменим $\frac{a}{b}$ через x; неравенство примет вид

$$(1+x)^p\leqslant 2^{p-1}(1+x^p),$$

или

$$2^{p-1}(1+x^p)-(1+x)^p\geqslant 0.$$

---

[1]) Например, у функции $y=x^3$ производная $y'=3x^2$ равна нулю при $x=0$, хотя сама функция возрастает.

Положим

$$2^{p-1}(1+x^p)-(1+x)^p=f(x);$$

очевидно, что

$$f'(x)=2^{p-1}\,px^{p-1}-p\,(1+x)^{p-1}=p\,[(2x)^{p-1}-(1+x)^{p-1}]<0$$

при $0<x<1$ и $p>1$ (так как $2x<1+x$ и $(2x)^{p-1}<(1+x)^{p-1}$). Поэтому $f(x)$ убывает при возрастании x и, следовательно,

$$f(x)>f(1) \quad \text{при} \quad 0<x<1.$$

Замечая, что $f(1)=2^{p-1}\cdot 2-2^p=0$, получим:

$$f(x)>0 \quad \text{при} \quad 0<x<1;$$

если $x=1$, то это неравенство обращается в равенство. Заменяя $f(x)$ её выражением, а затем полагая $x=\frac{a}{b}$, получаем последовательно:

$$2^{p-1}(1+x^p)-(1+x)^p\geqslant 0 \quad \text{и} \quad (a+b)^p\leqslant 2^{p-1}(a^p+b^p).$$

Так как это неравенство симметрично относительно a и b, то к тому же результату мы придём и тогда, когда $b\leqslant a$.

Предоставляем читателю доказать, что при $p\leqslant 1$ и любых положительных a и b справедливо неравенство:

$$(a+b)^p\leqslant a^p+b^p$$

(например, $\sqrt[n]{a+b}\leqslant\sqrt[n]{a}+\sqrt[n]{b}$ при любом натуральном n и $a>0$, $b>0$).

Пример 3. Докажем следующую теорему Гюйгенса: если p_n и P_n — периметры правильных n-угольников, вписанного в окружность радиуса R и описанного около неё, то

$$\frac{2}{3}p_n+\frac{1}{3}P_n>2\pi R.$$

Заметим сначала, что

$$p_n=2Rn\sin\frac{\pi}{n} \quad \text{и} \quad P_n=2Rn\operatorname{tg}\frac{\pi}{n};$$

поэтому

$$\frac{2}{3}p_n+\frac{1}{3}P_n=\frac{Rn}{3}\left(2\sin\frac{\pi}{n}+\operatorname{tg}\frac{\pi}{n}\right)$$

и неравенство принимает вид

$$2\sin\frac{\pi}{n}+\operatorname{tg}\frac{\pi}{n}>3\frac{\pi}{n}.$$

Установим более общее неравенство

$$2\sin x+\operatorname{tg}x>3x, \quad 0<x<\frac{\pi}{2}, \tag{1}$$

откуда и будет следовать теорема Гюйгенса.

Рассмотрим функцию

$$f(x)=2\sin x+\operatorname{tg}x-3x;$$

её производная имеет вид

$$f'(x)=2\cos x+\frac{1}{\cos^2 x}-3=\frac{2\cos^3 x-3\cos^2 x+1}{\cos^2 x}=$$
$$=\frac{(2\cos^3 x-2\cos^2 x)+(1-\cos^2 x)}{\cos^2 x}=\frac{(1-\cos x)(1+\cos x-2\cos^2 x)}{\cos^2 x}.$$

Здесь $1-\cos x>0$, $\cos^2 x>0$ и $1+\cos x-2\cos^2 x=1-\cos^2 x+$ $+\cos x(1-\cos x)>0$ при $0<x<\frac{\pi}{2}$; поэтому $f'(x)>0$, т. е. $f(x)$ возрастает, когда x возрастает от 0 до $\frac{\pi}{2}$ и, следовательно,

$$f(x)>0 \quad \text{при} \quad 0<x<\frac{\pi}{2}.$$

Это и есть требуемое неравенство.

Из неравенства (1), в частности, вытекает:

$$\sin x+\operatorname{tg}x>3x-\sin x>3x-x=2x,$$

т. е.

$$\sin x+\operatorname{tg}x>2x, \quad 0<x<\frac{\pi}{2}.$$

Приём, разъяснённый в рассмотренных трёх примерах, имеет общий характер; с его помощью можно получить многие другие неравенства.

15. Экстремум функции. Вообще говоря, те функции, с которыми приходится иметь дело на практике, не являются монотонными во всей области своего существования. Обычно графики их имеют вид вроде изображённого на рис. 13.

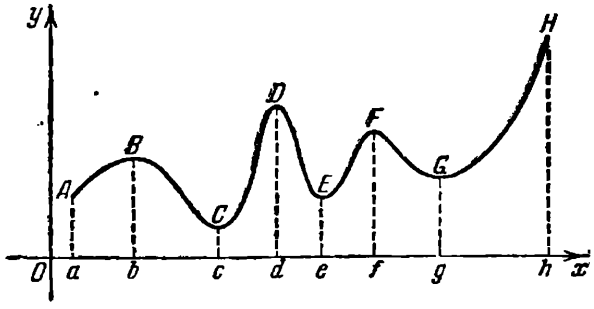


Рис. 13.

Рассматривая график рис. 13, мы видим на нём ряд характерных точек B, C, D, E, F, G. Значение функции, изображаемое каждой из этих точек, не будучи наибольшим или наименьшим среди всех значений её на промежутке $[a, h]$, является всё же таковым по сравнению со всеми значениями её для достаточно близких значений аргумента. Так, например, значение функции для $x=b$, изображаемое точкой B (или ординатой bB), является наибольшим на промежутке $[a, c]$. В связи с этими наблюдениями дадим

Определение. Говорят, что *функция* $f(x)$ *имеет при* $x=x_0$ *максимум (минимум), если существует промежуток* $[p, q]$, *содержащий точку* x_0 *внутри себя* (т. е. $p<x_0<q$) *и сам содержащийся в области задания функции, такой, что для всех* x *из* $[p, q]$ *оказывается* $f(x)\leqslant f(x_0)$ $[f(x)\geqslant f(x_0)]$.

Объединяющим термином для максимума и минимума служит термин *экстремум* [1]).

Полезно подчеркнуть, что по самому определению та точка, в которой функция имеет экстремум, должна лежать *внутри* промежутка задания функции, а не на его конце. Поэтому про функцию, изображённую на рис. 13, нельзя сказать, что она имеет максимум при $x=h$. Такое ограничение включено в определение экстремума для того, чтобы к точкам, где есть экстремум, можно было применить теорему Ферма (п° 10). Нетрудно видеть, что эта теорема допускает такую формулировку:

Теорема Ферма. *Пусть функция* $f(x)$ *в точке* $x=x_0$ *имеет экстремум. Если в этой точке существует производная* $f'(x_0)$, *то необходимо будет* $f'(x_0)=0$.

Само собой разумеется, что обратное заключение было бы неверно: из факта обращения производной в какой-либо точке в нуль совершенно не следует, что в этой точке есть экстремум. Например, производная функции $y=x^3$, равная $y'=3x^2$, обращается в нуль при $x=0$, но функция $y=x^3$ монотонна на всей числовой оси.

Те точки, в которых производная какой-либо функции обращается в нуль, называются *стационарными* точками. Геометрически это суть абсциссы тех точек графика функции, в которых касательная параллельна оси Ox.

Вышеприведённая форма теоремы Ферма означает, что точки, где есть экстремум, обязательно должны быть стационарными. Однако при этом заранее надо предположить существование производной в исследуемой точке. Например, функция $y=|x|$ (см. рис. 6) имеет минимум при $x=0$, но это — не стационарная точка, ибо в ней не существует производной.

Для изучения функций весьма важно уметь находить точки экстремума. В своей общей постановке задача эта весьма трудна, и мы рассмотрим её лишь для одного частного класса функций.

Будем говорить, что функция $f(x)$ входит в класс $K([a, b])$, если она обладает следующими свойствами:

1) Функция $f(x)$ задана на $[a, b]$ и во всех точках этого промежутка имеет производную [2]) $f'(x)$.

2) Производная $f'(x)$ непрерывна на $[a, b]$.

---

[1]) Слова «максимум», «минимум» и «экстремум» по-латыни означают соответственно «большее», «меньшее» и «крайнее» (подразумевается «значение функции»).

[2]) Отсюда уже вытекает непрерывность функции $f(x)$.

3) На (a, b) имеется разве лишь конечное число стационарных точек.

Для функций этого класса вопрос об отыскании экстремумов сравнительно прост. В самом деле, если стационарных точек на (a, b) вовсе нет, то $f(x)$ монотонна на $[a, b]$. Действительно, если бы в двух точках x' и x'' промежутка (a, b) производная $f'(x)$ имела разные знаки, то, будучи непрерывной, она должна была бы между этими точками обратиться в нуль, что, однако, невозможно ввиду отсутствия стационарных точек. Стало быть, во всех точках (a, b) производная $f'(x)$ имеет один и тот же знак, а тогда $f(x)$ монотонна.

Если же у $f(x)$ имеются на (a, b) стационарные точки $x_1 < x_2 < \ldots < x_m$ (причём это *все* без исключения её стационарные точки), то такое же рассуждение показывает, что на открытых промежутках (a, x_1), (x_1, x_2), $\ldots$, (x_m, b) производная $f'(x)$ сохраняет знак и $f(x)$ монотонна на замкнутых промежутках $[a, x_1]$, $[x_1, x_2]$, $\ldots$, $[x_m, b]$.

Сказанное здесь приводит к следующему правилу:

Правило. *Для нахождения экстремумов функции $f(x)$ класса $K([a, b])$ нужно*

1) *найти производную $f'(x)$,*

2) *положить $f'(x) = 0$ и решить полученное уравнение,*

3) *корни $x_1, x_2, \ldots, x_m$ предыдущего уравнения, лежащие в открытом промежутке (a, b), надо подвергнуть исследованию, определив знаки производной $f'(x)$ в каждом из промежутков (a, x_1), (x_1, x_2), $\ldots$, (x_m, b).*

Заключение о характере точки x_i делается по следующей схеме [1]):

	Знак $f'(x)$ при $x_{i-1} < x < x_i$	Знак $f'(x)$ при $x_i < x < x_{i+1}$	Заключение
1	+	+	Экстремума нет
2	+	−	Максимум
3	−	−	Экстремума нет
4	−	+	Минимум

В случаях 1 и 3 говорят, что точка $x = x_i$ есть точка перегиба.

Исследование функции на экстремум рекомендуется сопровождать построением её графика. На этом графике, помимо стационарных точек, следует отметить также и точки пересечения с осями. Если изучаемая функция задана на всей оси, то следует на графике показать, каково её поведение при безграничном возрастании и убывании аргумента.

---

[1]) Мы полагаем $x_0 = a$, $x_{m+1} = b$.

В тех случаях, когда в стационарной точке x_i существует вторая производная $f''(x_i)$, бывает удобно пользоваться следующим предложением:

Теорема. *Если x_i — стационарная точка, лежащая внутри промежутка $[a, b]$, где задана функция, и $f''(x_i) \neq 0$, то в точке x_i есть экстремум. Это — максимум при $f''(x_i) < 0$ и минимум при $f''(x_i) > 0$.*

В самом деле, полагая $x_i + \Delta x = x$, получим:

$$f''(x_i) = \lim_{x \to x_i} \frac{f'(x) - f'(x_i)}{x - x_i}.$$

Но $f'(x_i) = 0$, ибо x_i — точка стационарная. Значит,

$$f''(x_i) = \lim_{x \to x_i} \frac{f'(x)}{x - x_i}.$$

Пусть для определённости $f''(x_i) > 0$. Тогда для достаточно малых $|x - x_i|$ дробь

$$\frac{f'(x)}{x - x_i}$$

также будет положительна. Иначе говоря, для x, достаточно близких к x_i, знак числителя $f'(x)$ будет совпадать со знаком знаменателя $x - x_i$. Но тогда производная $f'(x)$ при переходе x из промежутка (x_{i-1}, x_i) в промежуток (x_i, x_{i+1}) будет менять знак с — на +, а это, как мы уже знаем, обеспечивает при $x = x_i$ наличие минимума.

Пример 1. Исследовать функцию

$$y = \frac{x}{1 + x^2}.$$

Здесь

$$y' = \frac{1 - x^2}{(1 + x^2)^2}.$$

Приравняв y' нулю, находим стационарные точки $x_1 = -1$ и $x_2 = +1$. Определяя знаки y' в промежутках $(-\infty, -1)$, $(-1, +1)$

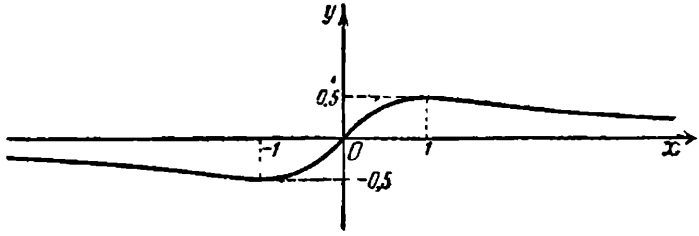


Рис. 14.

и $(+1, +\infty)$, видим, что знаки эти таковы: —, +, —. Значит, при $x = -1$ будет минимум, а при $x = +1$ — максимум. Замечая,

что $f(-1)=-\frac{1}{2}$, $f(+1)=+\frac{1}{2}$, что график функции проходит через начало координат и что как при $x\to+\infty$, так и при $x\to-\infty$ будет $\lim f(x)=0$, вычерчиваем график функции (рис. 14).

Пример 2. Исследовать функцию

$$y=x^3(x-5)^2.$$

Здесь $y'=5x^2(x-5)(x-3)$. Значит, стационарные точки $x_1=0$, $x_2=3$, $x_3=5$. Знаки y' в промежутках $(-\infty, 0)$, $(0, 3)$, $(3, 5)$ и $(5, +\infty)$ таковы: $+$, $+$, $-$, $+$. Значит, $x_1=0$ есть точка перегиба, $x_2=3$ — точка максимума и $x_3=5$ — точка минимума. Кроме того, $f(0)=0$, $f(3)=108$, $f(5)=0$. Ясно также, что график функции пересекает ось Ox при $x=0$ и касается её при $x=5$ и что

$$\lim_{x\to-\infty} y=-\infty, \quad \lim_{x\to+\infty} y=+\infty.$$

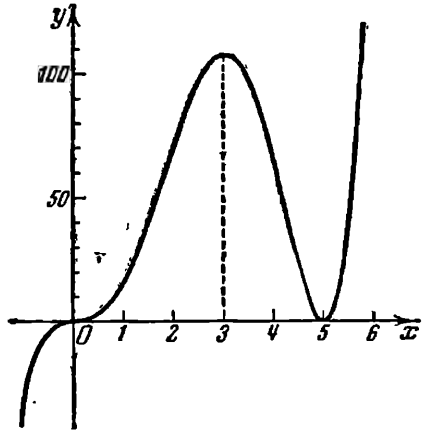


Рис. 15.

График функции изображён на рис. 15 [1]).

Предлагаем читателю, пользуясь изложенными соображениями, разобрать те примеры ($y=x^3-x$ и др.), которые в предыдущей статье (см. стр. 49, 51 и др.) были рассмотрены без привлечения методов дифференциального исчисления.

16. Нахождение наибольшего и наименьшего значений функции на замкнутом промежутке. Пусть функция $f(x)$ принадлежит классу $K([a, b])$. Тогда она непрерывна и, как известно (см. стр. 218), имеет на $[a, b]$ наибольшее значение M и наименьшее значение m. Займёмся вопросом отыскания этих значений, остановившись для определённости на M. Пусть своё наибольшее значение M функция $f(x)$ принимает в точке x_0, $f(x_0)=M$.

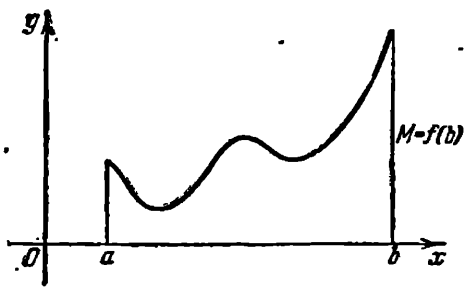


Рис. 16.

Если $a<x_0<b$, то в точке x_0 у функции $f(x)$, очевидно, будет *максимум*. Однако не исключено, что $x_0=a$ или $x_0=b$

[1]) При вычерчивании графика функции $y=x^3(x-5)^2$ мы выбрали различные единицы масштаба по оси Ox и по оси Oy. Иначе точка A, ордината которой равна 108, не поместилась бы на рис. 15.

(рис. 16), и тогда x_0 может даже и не быть стационарной точкой. Из сказанного вытекает следующее

Правило. *Чтобы найти наибольшее значение функции* $f(x)$ *на замкнутом промежутке* $[a, b]$, *нужно найти все её точки максимума, лежащие на открытом промежутке* (a, b). *Если это суть точки* $x_1, x_2, \ldots, x_n$, *то искомым наибольшим значением будет наибольшее из конечного множества чисел*

$$f(x_1),\ f(x_2),\ \ldots,\ f(x_n),\ f(a),\ f(b).$$

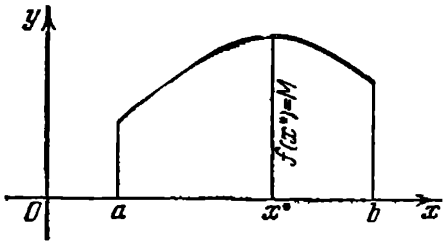


Рис. 17.

Отметим один важный частный случай: *если в открытом промежутке* (a, b) *имеется только одна точка экстремума* x^* *и это — точка максимума, то* $f(x^*)$ *и будет наибольшим значением* $f(x)$ на $[a, b]$.

Справедливость этого утверждения ясна из рис. 17.

Приведём две задачи конкретного характера, решающиеся при помощи изложенной теории.

Задача 1. Имеется прямоугольный лист жести размером $8\,дм \times 5\,дм$ (рис. 18). Требуется вырезать по углам листа такие одинаковые квадратики, чтобы после загибания оставшихся кромок получилась открытая сверху коробка наибольшего объёма.

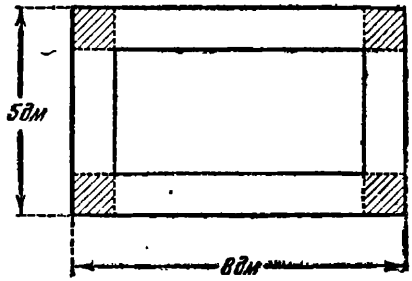


Рис. 18.

Обозначим через x сторону вырезаемого квадрата. Тогда $0 \leqslant x \leqslant 2{,}5$. Очевидно, объём коробки, упоминаемой в условии задачи, таков:

$$V = x(8-2x)(5-2x) = 4x^3 - 26x^2 + 40x.$$

Дело свелось к нахождению наибольшего значения этой функции на промежутке $\left[0,\ 2\frac{1}{2}\right]$.

Дифференцируя дважды, находим

$$V' = 12x^2 - 52x + 40, \quad V'' = 24x - 52.$$

Корни уравнения $V'=0$ суть $x_1 = 1, x_2 = 3\frac{1}{3}$. Из них в открытом промежутке $\left(0, 2\frac{1}{2}\right)$ лежит только $x_1 = 1$. Так как $V''(1) = -28 < 0$, то при $x = 1$ имеется максимум, и, как указано выше, здесь достигается и искомое наибольшее значение. Итак, сторона искомого квадрата должна равняться 1 *дм*.

Задача 2. В данный конус вписать цилиндр наибольшего объёма (рис. 19).

Обозначая радиус и высоту цилиндра через r и h, имеем:

$$V=\pi r^2 h.$$

С другой стороны, из подобия треугольников ABC и MNC находим:

$$\frac{h}{H}=\frac{R-r}{R},$$

где R и H — радиус и высота конуса. Отсюда

$$V=\pi\frac{H}{R}(Rr^2-r^3).$$

Дело сводится к нахождению наибольшего значения этой функции (аргумента r) в промежутке $[0, R]$.

Дифференцируя дважды, находим:

$$V'=\pi\frac{H}{R}(2Rr-3r^2),$$

$$V''=\pi\frac{H}{R}(2R-6r).$$

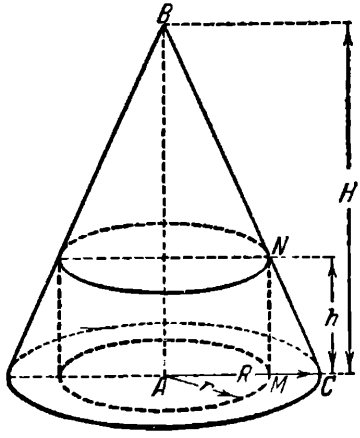


Рис. 19.

Корни уравнения $V'=0$ суть $r=0$ и $r=\frac{2}{3}R$. Из них в открытом промежутке $(0, R)$ лежит лишь $\frac{2}{3}R$. Это — точка максимума, ибо $V''\left(\frac{2}{3}R\right)=-2\pi H<0$. Значит, радиус искомого цилиндра есть $r=\frac{2}{3}R$. Отсюда легко найти и остальные его элементы.

Мы ограничимся двумя приведёнными примерами, ибо на них с достаточной ясностью видна чрезвычайная сила методов дифференциального исчисления [1]).

Заметим ещё, что существуют разнообразные частные приёмы, позволяющие упростить процесс нахождения наибольшего и наименьшего значений функции. Аналогичная теория разработана также и для функций многих переменных, но обо всём этом мы здесь уже говорить не будем, отсылая за подробностями к специальным руководствам.

---

[1]) Во многих случаях, используя индивидуальные особенности задачи на нахождение наибольшего значения функции, подобную задачу удаётся решить и элементарными приёмами. Сила методов дифференциального исчисления состоит именно в возможности игнорировать индивидуальные особенности задачи — это *методы общие*. Положение вещей таково же, как при решении арифметических задач с помощью уравнений: подчас можно обойтись и без них, но уравнения доставляют *общий метод* решения арифметических задач.

ГЛАВА II

ИНТЕГРАЛЫ

§ 4. Неопределённые интегралы

17. Основные понятия. В дифференциальном исчислении основной операцией является нахождение производной заданной функции. Мы уже знаем, что существо дела здесь заключается в установлении скорости изменения этой функции по сравнению с аргументом. Весьма часто, однако, приходится решать обратную задачу, когда по заданной скорости течения какого-либо процесса требуется восстановить сам этот процесс. В этом случае с математической точки зрения вопрос приводится к отысканию функции по её производной. Эта операция, называемая *интегрированием*, является основной во второй половине математического анализа — интегральном исчислении.

Перейдём к точным определениям.

Определение. *Пусть функция* $f(x)$, *заданная в некотором промежутке* [1]) $[a, b]$, *во всех его точках является производной функции* $F(x)$, *также заданной в* $[a, b]$. *Тогда эта последняя функция* $F(x)$ *называется первообразной функцией для функции* $f(x)$ (в промежутке $[a, b]$).

Имеет место замечательная

Теорема 1. *У всякой непрерывной на промежутке* $[a, b]$ *функции имеется первообразная.*

Доказательство этой теоремы будет дано ниже в п° 24.

Нетрудно видеть, что, если функция $F(x)$ есть первообразная для $f(x)$, то функция $F(x)+C$ при любом постоянном C также является первообразной для $f(x)$. В то же время никаких других первообразных, кроме функций вида $F(x)+C$, у $f(x)$ уже быть не может. Действительно, если $F_1(x)$ есть какая-то первообразная для $f(x)$, то производная разности $F_1(x)-F(x)$ будет всюду на $[a, b]$ равняться нулю, а тогда, как было доказано в п° 14, сама разность есть величина постоянная, т. е.

$$F_1(x)-F(x)=C \quad \text{и} \quad F_1(x)=F(x)+C.$$

[1]) Этот промежуток может быть замкнутым, открытым или полуоткрытым. В тексте мы употребили обозначение замкнутого промежутка лишь для определённости.

Определение. *Если $F(x)$ есть первообразная функция для $f(x)$, то функция двух аргументов x и C, равная $F(x)+C$, называется неопределённым интегралом функции $f(x)$ и обозначается символом*

$$\int f(x)\,dx.$$

Таким образом, неопределённый интеграл какой-нибудь функции представляет собой общий вид первообразных функций для этой функции. Величина C, входящая в определение неопределённого интеграла, называется «произвольной постоянной». Придавая ей то или иное закреплённое значение, мы можем получить из неопределённого интеграла любую первообразную.

Легко понять, что из самого определения понятия интеграла вытекает следующее утверждение:

Теорема 2. *Производная неопределённого интеграла равна подинтегральной функции, т. е.*

$$\left(\int f(x)\,dx\right)' = f(x).$$

Для успешного применения интегрального исчисления нужна разработанная техника нахождения неопределённых интегралов от элементарных функций. Читатель, желающий приобрести таковую, должен обратиться к специальным руководствам. Мы же здесь дадим лишь общее представление об этом вопросе.

В основе упомянутой техники лежит некоторое количество простых формул. Мы ограничимся следующими формулами:

1. $\int dx = x + C.$
2. $\int x^a dx = \frac{x^{a+1}}{a+1} + C \quad (a \neq -1).$
3. $\int \frac{dx}{x} = \ln x + C.$ 4. $\int a^x dx = \frac{a^x}{\ln a} + C.$
5. $\int e^x dx = e^x + C.$ 6. $\int \cos x\,dx = \sin x + C.$
7. $\int \sin x\,dx = -\cos x + C.$ 8. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C.$
9. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C.$
10. $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C \quad (a > 0).$
11. $\int \frac{dx}{a^2 + x^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C \quad (a \neq 0).$
12. $\int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \frac{x-a}{x+a} + C \quad (a \neq 0).$

Проверка справедливости каждой из этих формул проводится при помощи дифференцирования её правой части, ибо из самого определения интеграла вытекает

Правило. *Чтобы установить справедливость равенства*

$$\int f(x)\,dx = F(x) + C, \qquad (*)$$

надо продифференцировать его правую часть. Если при этом получится подинтегральная функция части левой, то равенство (*) *верно.*

Мы не будем проводить проверки формул вышеприведённой таблицы, а лишь отметим те промежутки изменения x, в которых справедлива та или иная из этих формул.

Формулы 1, 4, 5, 6, 7, 11 справедливы на всей оси.

Формула 2 справедлива в тех промежутках, в которых имеют смысл обе её части. Например формула

$$\int \frac{dx}{\sqrt{x}} = 2\sqrt{x} + C,$$

верна в $(0, +\infty)$, формула

$$\int \sqrt[3]{x}\,dx = \frac{3}{4}\sqrt[3]{x^4} + C$$

верна в $(-\infty, +\infty)$ и т. п.

Формула 3 верна в $(0, +\infty)$. Если же $x \in (-\infty, 0)$, то вместо формулы 3 надо написать

$$\int \frac{dx}{x} = \ln(-x) + C.$$

Формула 8 верна в любом промежутке, не содержащем точек вида $(2n+1)\frac{\pi}{2}$, а формула 9 — в любом промежутке, не содержащем точек $n\pi$.

Формула 10 верна в $(-a, +a)$. Наконец, формула 12 верна в каждом из промежутков $(-\infty, -|a|)$ и $(|a|, +\infty)$.

Интегралы от более сложных элементарных функций стараются свести к вышеприведённым «табличным интегралам». Для этого существует целый ряд разнообразных приёмов. Простейшие из этих приёмов состоят в применении двух следующих предложений:

Теорема 3. *Интеграл от алгебраической суммы нескольких функций равен алгебраической сумме интегралов слагаемых, т. е.*

$$\int [f(x) + g(x) - h(x)]\,dx = \int f(x)\,dx + \int g(x)\,dx - \int h(x)\,dx. \qquad (*^*)$$

Теорема 4. *Постоянный множитель можно выносить за знак интеграла, т. е.*

$$\int a f(x)\,dx = a\int f(x)\,dx.$$

Проверим хотя бы формулу (*). Согласно правилу, высказанному выше, для этого надо продифференцировать правую часть этой формулы. Так как производная суммы равна сумме производных слагаемых, то надо дифференцировать по отдельности слагаемые правой части формулы (*). Применяя теорему 2, легко убеждаемся, что искомая производная есть $f(x)+g(x)-h(x)$, т. е. совпадает с подинтегральной функцией левой части. Аналогично доказывается и теорема 4.

18. Интегрирование с помощью подстановки. Чрезвычайно сильным методом приведения интеграла к табличной форме является метод подстановки или замены переменной. Он применяется в двух различных формах, каждая из которых основана на следующей теореме:

Теорема. Пусть $F(z)$ *есть на каком-нибудь промежутке* $[p, q]$ *первообразная функция для функции* $f(z)$. *Если* $\varphi(x)$ *есть дифференцируемая функция, заданная на промежутке* $[a, b]$ *и удовлетворяющая неравенствам* $p \leqslant \varphi(x) \leqslant q$, *то сложная функция* $F[\varphi(x)]$ *будет первообразной для функции* $f[\varphi(x)]\varphi'(x)$.

В самом деле, дифференцируя сложную функцию $y=F[\varphi(x)]$, мы должны ввести промежуточный аргумент $z=\varphi(x)$. Тогда $y=F(z)$, $z=\varphi(x)$ и $y'_x=y'_z\cdot z'_x=F'(z)\varphi'(x)$. Так как $F'(z)=f(z)$, то $y'_x=f(z)\varphi'(x)=f[\varphi(x)]\varphi'(x)$, чем и доказана теорема.

Доказанную теорему, очевидно, можно формулировать и так: если

$$\int f(z)\,dz = F(z)+C,$$

то

$$\int f[\varphi(x)]\varphi'(x)\,dx = F[\varphi(x)]+C.$$

Отсюда следует

Первое правило подстановки. *Чтобы вычислить интеграл*

$$I=\int f[\varphi(x)]\varphi'(x)\,dx,$$

записываем его в форме

$$I=\int f[\varphi(x)]\,d\varphi(x),$$

заменяем здесь $\varphi(x)$ *на* z, *вычисляем полученный интеграл и в найденном ответе производим обратную замену* z *на* $\varphi(x)$.

Приведём несколько примеров:

1) $\int \frac{dx}{x(1+\ln^2 x)} = \int \frac{d\ln x}{1+\ln^2 x} = \int \frac{dz}{1+z^2} = \operatorname{arctg} z + C = \operatorname{arctg}\ln x + C;$

2) $\int \frac{e^x dx}{1+e^{2x}} = \int \frac{de^x}{1+e^{2x}} = \int \frac{dz}{1+z^2} = \operatorname{arctg} z + C = \operatorname{arctg} e^x + C;$

3) $\int \frac{\cos x\, dx}{1+\sin^2 x} = \int \frac{d\sin x}{1+\sin^2 x} = \int \frac{dz}{1+z^2} = \operatorname{arctg} z + C = \operatorname{arctg}\sin x + C.$

Мы видим, что одна и та же табличная формула

$$\int \frac{dz}{1+z^2} = \operatorname{arctg} z + C$$

позволяет находить бесчисленное множество разнообразных интегралов.

Точно так же из формулы

$$\int \frac{dz}{z} = \ln z + C$$

следует, что

$$\int \operatorname{tg} x\, dx = \int -\frac{d\cos x}{\cos x} = -\ln\cos x + C,$$

$$\int \frac{x^2 dx}{x^3+2} = \frac{1}{3}\int \frac{d(x^3+2)}{x^3+2} = \frac{1}{3}\ln(x^3+2) + C,$$

$$\int \frac{dx}{(1+x^2)\operatorname{arctg} x} = \int \frac{d\operatorname{arctg} x}{\operatorname{arctg} x} = \ln\operatorname{arctg} x + C$$

и т. д.

Остановимся теперь на другой форме замены переменной. Пусть $f(x)$ — непрерывная функция, заданная на каком-нибудь промежутке $[a, b]$. Допустим, что $x = \varphi(t)$ есть функция, заданная на другом промежутке $[\alpha, \beta]$, имеющая там производную $\varphi'(t)$ и удовлетворяющая неравенствам $a \leqslant \varphi(t) \leqslant b$. Пусть, кроме того, существует обратная функция $t = \psi(x)$, заданная на $[a, b]$. Рассмотрим интеграл

$$I_1 = \int f[\varphi(t)]\,\varphi'(t)\,dt.$$

Согласно сказанному выше для нахождения этого интеграла нужно переписать его в форме

$$\int f[\varphi(t)]\,d\varphi(t)$$

и заменить $\varphi(t)$ через x, что приведёт нас к интегралу

$$I = \int f(x)\,dx.$$

Этот последний интеграл заведомо существует (ибо $f(x)$, будучи непрерывной, имеет первообразную). Пусть

$$\int f(x)\,dx = A(x) + C.$$

Тогда, применяя 1-е правило подстановки к интегралу I_1, мы получим:

$$I_1 = A[\varphi(t)] + C.$$

Пусть $A[\varphi(t)] = F(t)$. Заменяя здесь t через $\psi(x)$ и замечая, что $\varphi[\psi(x)] = x$, находим:

$$A(x) = F[\psi(x)].$$

Отсюда вытекает, что

$$\int f(x)\,dx = F[\psi(x)] + C.$$

Если ещё заметить, что $F(t)$ есть не что иное, как $A[\varphi(t)]$, т. е. первообразная для $f[\varphi(t)]\varphi'(t)$, то мы сможем формулировать

Второе правило подстановки. *Чтобы вычислить интеграл*

$$I = \int f(x)\,dx,$$

полагаем $x = \varphi(t)$, *где* $\varphi(t)$ — *дифференцируемая функция, имеющая обратную функцию* $\psi(x)$. *Вычислив полученный интеграл, заменим в нём* t *через* $\psi(x)$, *что и приводит к значению искомого интеграла* I.

Приведём два примера.

Пример 1. Пусть $I = \int \sqrt{1-x^2}\,dx$. Положим $x = \sin t$. Это приводит к интегралу

$$\int \cos^2 t\,dt = \int \frac{1+\cos 2t}{2}\,dt = \frac{t}{2} + \frac{1}{4}\int \cos 2t\,d(2t) = \frac{t}{2} + \frac{\sin 2t}{4} + C.$$

Значит,

$$I = \frac{\arcsin x}{2} + \frac{x\sqrt{1-x^2}}{2} + C.$$

Пример 2. $I = \int \frac{dx}{1+\sqrt{x}}$. Положим $x = t^2$. Тогда

$$I = \int \frac{2t\,dt}{1+t} = 2\int \left(1 - \frac{1}{1+t}\right) dt = 2t - 2\ln(1+t) + C =$$
$$= 2\sqrt{x} - 2\ln(1+\sqrt{x}) + C.$$

19. Интегрирование по частям. Другим довольно общим приёмом преобразования интеграла является так называемое «интегрирование по частям».

Пусть $u = u(x)$ и $v = v(x)$ суть две дифференцируемые функции, заданные на одном и том же промежутке $[a, b]$. Тогда на этом промежутке будет

$$(uv)' = u'v + uv'.$$

Последнее равенство можно переписать в равносильной форме

$$\int (u'v + uv')\,dx = uv + C.$$

Отсюда, замечая, что $u'dx = du$, $v'dx = dv$, получаем:

$$\int u\,dv = uv - \int v\,du, \qquad (1)$$

причём произвольная постоянная, находившаяся в правой части, включена в интеграл $\int v\,du$. Формула (1) называется *формулой интегрирования по частям*. Она представляет собой некое тождественное преобразование одного интеграла в другой. Если новый интеграл проще исходного, то формулу применять целесообразно.

Всматриваясь в строение формулы (1), мы замечаем, что для её применения к какому-либо интегралу надо подынтегральное выражение представить в форме произведения $u\,dv$ некоторой функции u на дифференциал другой функции dv. В результате же применения формулы (1) у нас появится интеграл от функции v, умноженной на дифференциал du. Иначе говоря, преобразование по формуле (1) состоит в интегрировании одного множителя dv и одновременном дифференцировании другого u. Вообще говоря, каждая из этих операций может привести к упрощению рассматриваемого интеграла, но чаще всё же это упрощение достигается за счёт дифференцирования множителя u. Поэтому некоторым указанием на целесообразность интегрирования по частям может служить наличие в составе подынтегральной функции такого множителя, который упрощается от дифференцирования. Этот множитель и следует принять за u, обозначив произведение остальных сомножителей подынтегрального выражения (включая dx!) через dv.

Обратимся к примерам.

1) $I = \int x^3 \ln x\,dx$.

Так как функция $\ln x$ упрощается от дифференцирования, то полагаем

$$\begin{array}{l|l} \ln x = u, & du = \dfrac{dx}{x}, \\ x^3 dx = dv, & v = \dfrac{x^4}{4}. \end{array}$$

Функция $v = \frac{x^4}{4}$ найдена нами с помощью интегрирования [1]) её дифференциала x^3dx. Имея в виду применить тождественное преобразование (1), мы не вводим при нахождении v произвольной постоянной, ибо при написании тождества (1) под v мы можем разуметь какую угодно определённую первообразную для dv. Это следует иметь в виду и в дальнейшем.

[1]) Этот дифференциал есть «часть» подынтегрального выражения. Отсюда и термин «интегрирование по частям».

Таким образом,

$$I=\frac{x^4}{4}\ln x-\int\frac{x^4}{4}\cdot\frac{dx}{x}=\frac{x^4}{4}\ln x-\frac{x^4}{16}+C.$$

2) $\int \operatorname{arctg} x\,dx=x\operatorname{arctg} x-\int\frac{x\,dx}{1+x^2}=x\operatorname{arctg} x-\frac{1}{2}\ln(1+x^2)+C.$

3) $\int x\cos x\,dx=\int x\,d\sin x=x\sin x-\int\sin x\,dx=x\sin x+\cos x+C.$

4) $\int xe^x dx=\int x\,de^x=xe^x-\int e^x dx=xe^x-e^x+C.$

20. Общие замечания по поводу интегрирования элементарных функций. Мы уже говорили выше, что у всякой непрерывной функции $f(x)$ имеется первообразная функция. Это обстоятельство следует сопоставить с тем, что существуют непрерывные функции, не имеющие производной.

Таким образом, если заниматься лишь вопросами существования у данной непрерывной функции первообразной и производной, то первый из этих вопросов решается всегда положительно, а второй — нет.

Иным окажется положение вещей, если мы будем рассматривать одни только элементарные функции и поставим вопрос о выражении их первообразных и производных снова через такие же функции. Именно, как мы уже знаем, всякая функция, являющаяся конечной комбинацией элементарных функций, не только обязательно имеет производную, но эта производная сама также есть конечная комбинация элементарных функций. По отношению к проблеме интегрирования дело обстоит совсем не так. Существуют очень простые элементарные функции, первообразные которых уже не выражаются никакой конечной комбинацией элементарных функций.

Так, например, можно доказать, что ни один из интегралов

$$\left.\begin{gathered}\int e^{-x^2}dx,\quad \int x\operatorname{tg} x\,dx,\quad \int\sqrt{\sin x}\,dx,\quad \int e^{\operatorname{arctg} x}dx,\\ \int\frac{e^x}{x}dx,\quad \int\frac{dx}{\ln x},\quad \int\frac{\sin x}{x}dx\end{gathered}\right\}\qquad(2)$$

не выражается конечным числом элементарных функций.

Чтобы разобраться в этом вопросе, следует прежде всего указать на то, что причисление какой-либо функции к классу «элементарных» функций есть вещь довольно условная. В конечном счёте ассортимент функций, которые в настоящее время принято называть элементарными, сложился в значительной степени под влиянием исторического хода развития математики не только как науки, но и (пожалуй, даже в большей степени) как учебного предмета.

Попробуем представить себе, как обстояло бы дело, если бы историческое развитие шло по-иному, и функция $\ln x$ не считалась

бы элементарной. Ясно, что и при этом воображаемом положении рациональную функцию $\frac{1}{x}$ всё же относили бы к разряду функций элементарных. И тогда интеграл

$$\int \frac{dx}{x} \tag{3}$$

оказался бы интегралом от элементарной функции, не выражающимся через элементарные.

Таким образом, мы видим, что проблема представления того или иного интеграла через элементарные функции получает точный смысл лишь при указании того, какие именно функции приняты за таковые. Расширяя запас элементарных функций, мы можем интеграл, не выражающийся через (старые) элементарные функции, превратить в выражающийся через (новые) элементарные функции. В приведённом только что примере достаточно ввести элементарную функцию $\ln x$, чтобы интеграл (3) выразился через неё.

Однако, встречаясь с тем или иным интегралом, не выражающимся через элементарные функции, не расширяют запаса элементарных функций. В самом деле, снова возвращаясь к нашему примеру, мы видим, что, не включая $\ln x$ в число элементарных функций, мы не будем и интеграл

$$\int \frac{dx}{\ln x} \tag{4}$$

считать интегралом от элементарной функции. Расширив запас элементарных функций за счёт введения функции $\ln x$, мы действительно будем в состоянии найти интеграл (3), но зато у нас появится интеграл (4), для выражения которого потребуется новое расширение класса элементарных функций. Таким образом, путь неограниченного расширения класса элементарных функций оказывается не приводящим к цели. В то же время сколько-нибудь значительное расширение класса элементарных функций связано с рядом неудобств. Тот исторически сложившийся класс элементарных функций, с которым имеет дело современный анализ, весьма удобен именно потому, что он не очень велик и многочисленные вычислительные формулы, связывающие функции этого класса, как, например, формулы

$$\sin^2 x + \cos^2 x = 1, \quad \sin(x+y) = \sin x \cos y + \sin y \cos x,$$

$$\log(xy) = \log x + \log y, \quad \frac{a^x}{a^y} = a^{x-y},$$

и т. п., легко запоминаются и известны уже выпускникам средней школы.

Надо заметить, однако, что многие интегралы, не выражающиеся через элементарные функции, весьма важны для решения самых раз-

нообразных вопросов. Так, например, интеграл $\int e^{-x^2}dx$ играет важную роль в теории вероятностей. В теории дифракции света встречаются интегралы

$$\int \sin x^2\,dx,\quad \int \cos x^2\,dx,$$

которые не могут быть выражены через элементарные функции. В исследованиях П. Л. Чебышева по распределению простых чисел существенное значение имеет интеграл $\int \frac{dx}{\ln x}$. Ниже (в п° 27) мы остановимся на вопросе об обращении с такими интегралами.

В свете высказанных соображений приобретает интерес вопрос о выделении таких частных классов элементарных функций, интегралы от которых также выражаются через элементарные функции. Приведём важную теорему, относящуюся к этому вопросу:

Теорема. *Интеграл от любой рациональной функции*

$$\frac{a_0x^n+a_1x^{n-1}+\ldots+a_n}{b_0x^m+b_1x^{m-1}+\ldots+b_m} \tag{5}$$

с действительными коэффициентами a_i и b_k выражается через рациональные функции, логарифмы и арктангенсы.

В самом деле, в алгебре устанавливается, что любая функция вида (5) (с действительными a_i и b_k) представима в форме суммы конечного числа слагаемых следующих 5 типов:

1) $A_0x^r+A_1x^{r-1}+\ldots+A_r$, 4) $\frac{Ax+B}{x^2+px+q}$,

2) $\frac{A}{x-a}$,

3) $\frac{A}{(x-a)^r}$ $(r>1)$, 5) $\frac{Ax+B}{(x^2+px+q)^r}$ $(r>1)$,

где все коэффициенты также действительны, а корни трёхчлена x^2+px+q — мнимые.

Что касается выражений 1), 2) и 3), то их интегрирование в элементарных функциях выполняется по формулам

$$\int\left(\sum_{k=0}^{r}A_kx^{r-k}\right)dx=\sum_{k=0}^{r}A_k\frac{x^{r-k+1}}{r-k+1}+C,$$

$$\int\frac{A}{x-a}\,dx=A\ln(x-a)+C,$$

$$\int\frac{A}{(x-a)^r}\,dx=\frac{A}{(1-r)(x-a)^{r-1}}+C.$$

Для интегрирования выражения типа 4) следует числитель его $Ax+B$ разделить на производную знаменателя, т. е. на $2x+p$. Представив $Ax+B$ в форме $\frac{A}{2}(2x+p)+\left(B-\frac{A}{2}p\right)$, будем иметь:

$$\int\frac{Ax+B}{x^2+px+q}\,dx=\frac{A}{2}\int\frac{2x+p}{x^2+px+q}\,dx+\left(B-\frac{A}{2}p\right)\int\frac{dx}{x^2+px+q}.$$

Первый из интегралов правой части равен $\ln(x^2+px+q)$, а для вычисления второго преобразуем его так:

$$\int \frac{dx}{x^2+px+q} = \int \frac{d\left(x+\frac{p}{2}\right)}{\left(x+\frac{p}{2}\right)^2+\left(q-\frac{p^2}{4}\right)}.$$

Замечая, что $q-\frac{p^2}{4}>0$ (ибо корни трёхчлена x^2+px+q — мнимые), мы видим, что последний интеграл равен [см. формулу 11) на стр. 367]

$$\frac{1}{\sqrt{q-\frac{p^2}{4}}}\operatorname{arctg}\frac{x+\frac{p}{2}}{\sqrt{q-\frac{p^2}{4}}}+C.$$

Итак, всё свелось к рассмотрению интеграла

$$I=\int \frac{Ax+B}{(x^2+px+q)^r}\,dx \qquad (r>1).$$

Исходя из уже отмеченного тождества

$$Ax+B=\frac{A}{2}(2x+p)+\left(B-\frac{A}{2}p\right),$$

получаем:

$$I=\frac{A}{2}\int \frac{(2x+p)\,dx}{(x^2+px+q)^r}+\left(B-\frac{A}{2}p\right)\int \frac{dx}{(x^2+px+q)^r}.$$

Первый интеграл правой части есть

$$\frac{A}{2(1-r)}\cdot\frac{1}{(x^2+px+q)^{r-1}}+C.$$

Второй же интеграл представляется так:

$$\int \frac{d\left(x+\frac{p}{2}\right)}{\left[\left(x+\frac{p}{2}\right)^2+\left(q-\frac{p^2}{4}\right)\right]^r}.$$

Полагая $x+\frac{p}{2}=z$, $q-\frac{p^2}{4}=a^2$, сведём дело к нахождению интеграла

$$I_r=\int \frac{dz}{(z^2+a^2)^r}.$$

Этот последний интеграл можно преобразовать к такому же интегралу, но с меньшим значком. Именно,

$$I_r=\frac{1}{a^2}\int \frac{z^2+a^2-z^2}{(z^2+a^2)^2}\,dz=\frac{1}{a^2}I_{r-1}-\frac{1}{a^2}\int \frac{z^2\,dz}{(z^2+a^2)^r}.$$

Полагая $u=z$, $dv=\frac{z\,dz}{(z^2+a^2)^r}$ $\left[\text{откуда } v=\frac{1}{2(1-r)}\cdot\frac{1}{(z^2+a^2)^{r-1}}\right]$ и интегрируя по частям, находим

$$\int \frac{z^2\,dz}{(z^2+a^2)^r}=\frac{1}{2(1-r)}\cdot\frac{z}{(z^2+a^2)^{r-1}}-\frac{1}{2(1-r)}\int\frac{dz}{(z^2+a^2)^{r-1}}=$$
$$=\frac{1}{2(1-r)}\cdot\frac{z}{(z^2+a^2)^{r-1}}-\frac{1}{2(1-r)}I_{r-1}.$$

Таким образом,

$$I_r = \frac{1}{a^2}\left[1 + \frac{1}{2(1-r)}\right] I_{r-1} - \frac{1}{2(1-r)} \cdot \frac{z}{(z^2+a^2)^{r-1}}.$$

Повторно применяя эту формулу, сведём дело к нахождению интеграла

$$I_1 = \int \frac{dz}{z^2+a^2} = \frac{1}{a} \operatorname{arctg} \frac{z}{a} + C.$$

Приведённый мелким шрифтом способ доказательства теоремы является в то же время и способом фактического вычисления интегралов от функций вида (5), но на практике обычно применяются более удобные способы интегрирования, наиболее важный из которых принадлежит М. В. Остроградскому. Мы не можем излагать этих способов, отсылая читателя к специальным руководствам.

Что касается иррациональных функций, то вопрос об их интегрировании в элементарных функциях гораздо более сложен. Укажем, например, на следующий результат по поводу интегрирования так называемых «биномиальных дифференциалов».

Теорема П. Л. Чебышева. *Интеграл*

$$\int x^m (ax^n + b)^p \, dx$$

при рациональных m, n и p выражается элементарно только в трёх случаях:

1) p — *целое*, 2) $\frac{m+1}{n}$ — *целое*, 3) $p + \frac{m+1}{n}$ — *целое*.

Мы не можем входить в дальнейшие подробности по затронутому вопросу, так же как и рассматривать ещё более сложный вопрос об элементарном выражении интегралов от трансцендентных функций.

§ 5. Определённые интегралы

21. Задачи, приводящие к понятию определённого интеграла. Основным понятием интегрального исчисления является всё же не понятие неопределённого интеграла, а понятие интеграла определённого. Оно существенно сложнее и целесообразно предпослать ему некоторые задачи конкретного характера, которые выясняют необходимость введения этого понятия.

I. ЗАДАЧА О МАССЕ СТЕРЖНЯ. Ещё в п° 1 мы ввели понятия средней плотности стержня и его истинной плотности в данной точке.

Если стержень однороден, то его истинная плотность одинакова во всех его точках и равна его средней плотности. У неоднородного же стержня истинная плотность p меняется от точки к точке. Если определять положение каждой точки M стержня с помощью

расстояния x её от одного из концов стержня (рис. 20), то его плотность p в точке x будет функцией от x, $p=p(x)$. Поставим задачу, как, зная эту функцию и длину l стержня, найти его массу m.

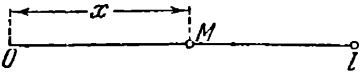


Рис. 20.

При решении этой задачи будем считать плотность $p(x)$ непрерывной функцией. Переходя к решению, разделим стержень точками $x_1<x_2<\ldots<x_{n-1}$ $(0<x_k<l)$ на n небольших участков (рис. 21). Для единообразия обозначений положим ещё $x_0=0$, $x_n=l$, и пусть λ есть наибольшая из разностей $x_{k+1}-x_k$. Отдельный участок $[x_k, x_{k+1}]$ стержня приближённо можно считать однородным [ибо из-за его малости (непрерывная) функция $p(x)$ не успевает на нём сколько-нибудь заметно измениться]. Делая такое допущение, мы тем самым принимаем плотность $p(x)$ на участке $[x_k, x_{k+1}]$ за постоянную. Пусть значение этой постоянной есть $p(\xi_k)$, где ξ_k есть произвольно выбранная точка участка $[x_k, x_{k+1}]$. Тогда масса участка $[x_k, x_{k+1}]$ будет равна $p(\xi_k)(x_{k+1}-x_k)$, а полная масса стержня будет

Рис. 21.

$$m=\sum_{k=0}^{n-1} p(\xi_k)(x_{k+1}-x_k).$$

Полученное выражение массы является, однако, лишь приближённым, ибо на самом деле отдельные участки стержня не однородны. Тем не менее, чем короче эти участки, т. е. чем меньше число λ, тем более точным будет найденное выражение m. Отсюда следует, что точное значение массы таково:

$$m=\lim_{\lambda\to 0}\sum_{k=0}^{n-1} p(\xi_k)(x_{k+1}-x_k). \qquad (1)$$

II. ЗАДАЧА О ПРОЙДЕННОМ ПУТИ. Пусть точка M движется по прямой, обладая скоростью v. Эта скорость меняется с течением времени и потому является функцией от времени t, $v=v(t)$. Поставим задачу — найти путь s, пройденный точкой за промежуток времени от момента $t=a$ до момента $t=b$.

При решении задачи будем считать скорость $v(t)$ непрерывной функцией t. Переходя к решению, разделим $[a, b]$ точками $t_1<t_2<\ldots<t_{n-1}$ $(a<t_k<b)$ на n коротких промежутков времени. Для единообразия положим ещё $t_0=a$, $t_n=b$ и пусть $\lambda=\max\{t_{k+1}-t_k\}$.

Так как за короткий промежуток времени $[t_k, t_{k+1}]$ скорость $v(t)$ (будучи непрерывной функцией) почти не меняется, то можно при-

ближённо считать её за этот промежуток времени постоянной и равной $v(\tau_k)$, где $\tau_k \in [t_k, t_{k+1}]$. С точки зрения механики это означает, что мы считаем движение точки за время $[t_k, t_{k+1}]$ равномерным. Но тогда путь, пройденный точкой за это время, очевидно, равен $v(\tau_k)(t_{k+1} - t_k)$, а путь, пройденный за всё время $[a, b]$, будет

$$s = \sum_{k=0}^{n-1} v(\tau_k)(t_{k+1} - t_k).$$

Полученное выражение для s, будучи лишь приближённым, оказывается тем более точным, чем меньше λ. Поэтому точное значение пути s таково:

$$s = \lim_{\lambda \to 0} \sum_{k=0}^{n-1} v(\tau_k)(t_{k+1} - t_k). \quad (2)$$

III. ЗАДАЧА О ПЛОЩАДИ КРИВОЛИНЕЙНОЙ ТРАПЕЦИИ. Рассмотрим плоскую фигуру, ограниченную линиями $y = 0$, $x = a$, $x = b$ и $y = f(x)$, где $f(x)$ есть непрерывная положительная функция, заданная при $a \leqslant x \leqslant b$ (рис. 22). Такая фигура называется *криволинейной трапецией*. Поставим вопрос о *площади* F этой трапеции.

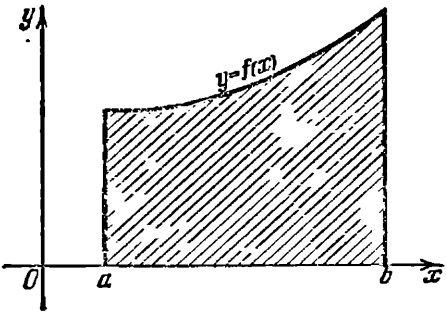


Рис. 22.

Отметим, что здесь, в отличие от двух рассмотренных выше задач, речь должна итти прежде всего о самом определении того, что́ такое площадь, и лишь затем — о нахождении её численного значения. Нижеприводимое рассуждение освещает оба эти момента.

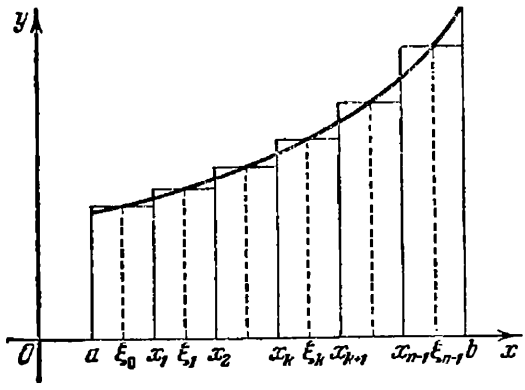


Рис. 23.

Разделим $[a, b]$ точками $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$ и пусть $\lambda = \max(x_{k+1} - x_k)$. Прямые $x = x_k$ разбивают нашу трапецию на n узких полос. Так как функция $f(x)$ непрерывна, то она мало меняется при $x_k \leqslant x \leqslant x_{k+1}$ и без большой погрешности её можно считать на промежутке $[x_k, x_{k+1}]$ постоянной и равной $f(\xi_k)$, где ξ_k есть произвольно взятая точка промежутка $[x_k, x_{k+1}]$. Легко видеть, что сделанное допущение равносильно тому, что мы принимаем выше-

упомянутые полосы за прямоугольники, а всю нашу трапецию — за ступенчатую фигуру, изображённую на рис. 23.

Площадь этой ступенчатой фигуры, очевидно, равна

$$F_{\text{ступ}} = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k).$$

Естественно считать, что эта площадь при малом λ является приближённым значением интересующей нас площади F. Поэтому мы по определению будем называть *площадью* нашей криволинейной трапеции предел

$$F = \lim_{\lambda \to 0} \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k), \tag{3}$$

причём, однако, здесь подлежит доказательству существование этого предела [1]).

Сравнивая выражения (1), (2) и (3), полученные в процессе решения рассмотренных задач, мы замечаем, что с чисто аналитической точки зрения все эти выражения совершенно одинаковы. Поэтому мы займёмся изучением этих выражений, называемых определёнными интегралами, уже не интересуясь их конкретным истолкованием.

22. Определённый интеграл. Дадим теперь точное определение понятия определённого интеграла.

Пусть на замкнутом промежутке $[a, b]$ задана функция $f(x)$. Проделаем следующие операции:

1) Раздробим $[a, b]$ на части точками

$$x_0 = a < x_1 < x_2 < \ldots < x_{n-1} < x_n = b,$$

причём наибольшую из разностей $x_{k+1} - x_k$ обозначим через λ.

2) В каждом частичном промежутке $[x_k, x_{k+1}]$ выберем по точке ξ_k и вычислим $f(\xi_k)$.

3) Умножим $f(\xi_k)$ на длину $(x_{k+1} - x_k)$ соответствующего промежутка $[x_k, x_{k+1}]$.

4) Сложим все найденные произведения. Сумму

$$\sigma = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k)$$

мы будем называть «интегральной суммой».

5) Будем изменять произведённое дробление $[a, b]$ так, чтобы величина λ стремилась к нулю.

---

[1]) В предыдущих двух случаях существование этого предела мы считали очевидным, ибо масса m и путь s — это заведомо существующие физические величины.

Если при этом существует конечный предел

$$I = \lim_{\lambda \to 0} \sigma, \tag{4}$$

не зависящий от выбора точек ξ_k, то этот предел называется *определённым интегралом* от функции $f(x)$ по промежутку $[a, b]$ и обозначается через

$$\int_a^b f(x)\,dx.$$

Точный смысл соотношения (4) таков: всякому $\varepsilon > 0$ отвечает такое $\delta > 0$, что при любом способе дробления, у которого $\lambda < \delta$, будет

$$|\sigma - I| < \varepsilon,$$

как бы при этом ни были выбраны точки $\xi_k \in [x_k, x_{k+1}]$.

Читатель видит, что предельное соотношение (4) имеет довольно своеобразный характер.

Отдадим себе отчёт в том, для каких функций введённое понятие оказывается достаточно естественным.

Если у функции $f(x)$ существует интеграл (в этом случае говорят, что $f(x)$ интегрируема), то это означает, что суммы σ, отвечающие дроблениям с достаточно малым λ, будут близки к некоторому постоянному числу, как бы ни выбирать точки ξ_k. Поэтому, меняя точки ξ_k, мы не будем существенно изменять величины суммы σ. Но это возможно лишь за счёт того, что изменение точек ξ_k не вызывает заметного изменения чисел $f(\xi_k)$ (по крайней мере в большинстве слагаемых суммы σ). Для функций непрерывных указанное обстоятельство и в самом деле имеет место, ибо точки ξ_k могут изменяться лишь в коротких промежутках $[x_k, x_{k+1}]$, а у непрерывных функций близким значениям аргумента отвечают близкие же значения функции. Поэтому естественно ожидать, что у непрерывной функции определённый интеграл существует. Если же функция $f(x)$ разрывна, то, вообще говоря, нет оснований ожидать у неё существования интеграла. Изложенные соображения подтверждаются следующей теоремой:

Теорема. *Если функция $f(x)$ непрерывна на $[a, b]$, то интеграл*

$$\int_a^b f(x)\,dx$$

существует.

Доказательство этой теоремы довольно сложно, и мы предпошлём ему некоторые вспомогательные соображения.

Пусть на $[a, b]$ задана непрерывная функция $f(x)$. Разобьём $[a, b]$ на части точками

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

и обозначим через M_k и m_k наибольшее и наименьшее значения $f(x)$ на частичном промежутке $[x_k, x_{k+1}]$. Суммы

$$S = \sum_{k=0}^{n-1} M_k (x_{k+1} - x_k), \quad s = \sum_{k=0}^{n-1} m_k (x_{k+1} - x_k)$$

называются соответственно *верхней* и *нижней интегральными суммами* (отвечающими выбранному способу дробления).

Введение этих сумм вызвано тем обстоятельством, что они полностью определяются способом дробления $[a, b]$, в то время как для определения суммы σ надо задать ещё точки ξ_k.

Легко видеть, что при выбранном способе дробления и при любом выборе точек ξ_k будет

$$s \leqslant \sigma \leqslant S. \tag{5}$$

Лемма. *Пусть промежуток $[a, b]$ раздроблён на части точками $x_0 = a < x_1 < x_2 < \ldots < x_n = b$, и составлены суммы S и s, отвечающие этому способу дробления. Если мы добавим новые точки дробления* (сохраняя старые) *и снова составим верхнюю и нижнюю суммы S' и s', то окажется*

$$s \leqslant s' \leqslant S' \leqslant S.$$

Иными словами, *от добавления новых точек деления нижняя интегральная сумма не уменьшается, а верхняя не увеличивается.*

Доказательство проведём лишь для верхних сумм. Очевидно, достаточно рассмотреть тот случай, когда вводится только одна новая точка деления $\bar{x}$, ибо общий случай приводится к этому путём повторного введения по одной точке.

Пусть $x_i < \bar{x} < x_{i+1}$. Тогда новая верхняя сумма S' получается из старой суммы S заменой слагаемого

$$M_i (x_{i+1} - x_i) \tag{6}$$

суммой двух слагаемых

$$M_i' (\bar{x} - x_i) + M_i'' (x_{i+1} - \bar{x}), \tag{7}$$

где M_i' и M_i'' суть наибольшие значения $f(x)$ в промежутках $[x_i, \bar{x}]$ и $[\bar{x}, x_{i+1}]$. Все же остальные слагаемые в обеих суммах S и S' одинаковы. Так как промежутки $[x_i, \bar{x}]$ и $[\bar{x}, x_{i+1}]$ суть части промежутка $[x_i, x_{i+1}]$, то, очевидно, $M_i' \leqslant M_i$, $M_i'' \leqslant M_i$. Но тогда

$$M_i' (\bar{x} - x_i) + M_i'' (x_{i+1} - \bar{x}) \leqslant M_i [(\bar{x} - x_i) + (x_{i+1} - \bar{x})] = M_i (x_{i+1} - x_i).$$

Иначе говоря, сумма (7) не больше величины (6), а тогда $S' \leqslant S$.

Следствие. *Каждая нижняя интегральная сумма не больше, чем каждая верхняя.*

В самом деле, выберем какие-либо два способа (I) и (II) дробления промежутка $[a, b]$ и пусть s_1 есть нижняя сумма, отвечающая способу (I), а S_2 — верхняя сумма, отвечающая способу (II). Составим новый способ дробления (III), точками деления которого служат точки деления обоих способов (I) и (II), и пусть нижняя и верхняя суммы, отвечающие этому новому способу дробления, суть s_3 и S_3. Согласно лемме будет $s_1 \leqslant s_3$, $S_3 \leqslant S_2$. С другой стороны, очевидно, $s_3 \leqslant S_3$. Отсюда и вытекает, что

$$s_1 \leqslant S_2.$$

Закрепим теперь какую-нибудь верхнюю сумму S_0. Тогда для каждой нижней суммы s будет

$$s \leqslant S_0.$$

Таким образом, множество нижних сумм $\{s\}$, отвечающих всевозможным способам дробления $[a, b]$, ограничено сверху, и S_0 — его верхняя граница. Обозначим через I *точную* верхнюю границу упомянутого множества

$$I = \sup \{s\}.$$

Тогда $I \leqslant S_0$, а так как S_0 есть произвольная верхняя сумма, то постоянное число I оказывается удовлетворяющим неравенству

$$s \leqslant I \leqslant S, \tag{8}$$

в котором s и S суть совершенно произвольные нижняя и верхняя суммы.

Теперь нетрудно доказать формулированную выше теорему (см. стр. 381). Пусть s, S и σ суть суммы, отвечающие какому-либо способу дробления (разумеется, для построения суммы σ надо указать ещё точки ξ_k). Из (5) и (8) следует, что

$$|\sigma - I| \leqslant S - s. \tag{9}$$

С другой стороны, функция $f(x)$, будучи непрерывной на замкнутом промежутке $[a, b]$, оказывается и равномерно непрерывной на этом промежутке. Значит, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что как только $|x'' - x'| < \delta$ (причём x' и x'' взяты из $[a, b]$), так сейчас же

$$|f(x'') - f(x')| < \frac{\varepsilon}{b-a}.$$

Предположим, что нами рассматривается такой способ дробления, у которого $\lambda < \delta$. Тогда, очевидно,

$$M_k - m_k < \frac{\varepsilon}{b-a} \qquad (k = 0, 1, 2, \ldots, n-1),$$

и потому

$$S - s = \sum_{k=0}^{n-1} (M_k - m_k)(x_{k+1} - x_k) < \frac{\varepsilon}{b-a} \sum_{k=0}^{n-1} (x_{k+1} - x_k) = \varepsilon.$$

Отсюда и из (9) следует, что

$$|\sigma - I| < \varepsilon. \tag{10}$$

Итак, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что при любом способе дробления, у которого $\lambda < \delta$, оказывается выполненным неравенство (10) (как бы ни выбирать точки ξ_k). Но это и означает, что

$$I = \lim_{\lambda \to 0} \sigma,$$

так что I и есть интеграл от функции $f(x)$. Теорема доказана.

Сопоставляя доказанную теорему с решением задачи III из п° 21, мы видим, что криволинейная трапеция, рассмотренная в упомянутой задаче, имеет площадь F, причём эта площадь выражается формулой

$$F = \int_a^b f(x)\,dx.$$

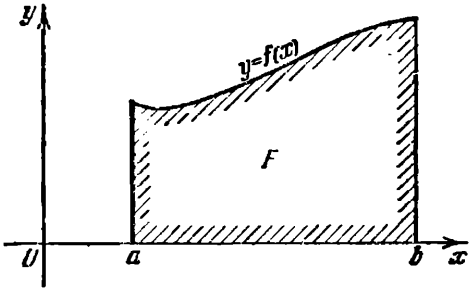


Рис. 24.

Читая эту формулу справа налево, находим

Геометрический смысл определённого интеграла. *Если $f(x)$ непрерывна и положительна на $[a, b]$, то интеграл*

$$\int_a^b f(x)\,dx$$

представляет собой площадь криволинейной трапеции, ограниченной линиями $y = 0$, $x = a$, $x = b$, $y = f(x)$ (рис. 24).

Не следует думать, что условие непрерывности функции необходимо для того, чтобы у неё существовал определённый интеграл. Интеграл может существовать и у разрывной функции. Пусть, например, функция $f(x)$, заданная на промежутке $[a, b]$, равна нулю

во всех точках этого промежутка, кроме конечного числа точек $z_1, z_2, \ldots, z_N$. Составим для $f(x)$ интегральную сумму σ.

Пусть из точек $\xi_0, \xi_1, \ldots, \xi_{n-1}$, входящих в определение σ, p точек совпадают с точками z_i, а остальные отличны от них. Тогда в сумме σ будет лишь p слагаемых, отличных от нуля. Если наибольшее из чисел $|f(z_i)|$ $(i=1, 2, \ldots, N)$ есть K, то, очевидно,

$$|\sigma| \leqslant Kp\lambda \leqslant KN\lambda,$$

откуда ясно, что при $\lambda \to 0$ будет и $\sigma \to 0$. Таким образом, интеграл

$$\int_a^b f(x)\,dx$$

существует и равен нулю.

Приведём теперь пример функции, не имеющей интеграла. Пусть $\varphi(x)$ задана на промежутке $[0, 1]$ так:

$$\varphi(x) = \begin{cases} 0, & \text{если } x \text{ — число иррациональное,} \\ 1, & \text{» } x \text{ — » рациональное.} \end{cases}$$

Если мы, составляя сумму σ, за точки ξ_k выберем числа иррациональные, то окажется $\sigma = 0$. Если же все ξ_k взять рациональными, то получится $\sigma = 1$. Таким образом, за счёт одного лишь уменьшения λ нельзя приблизить σ к какому-либо постоянному числу, и интеграл

$$\int_0^1 \varphi(x)\,dx$$

не существует.

В настоящее время известны точные признаки, позволяющие судить, имеет или нет заданная функция определённый интеграл, но мы ограничимся вышеприведённой теоремой об интегрируемости непрерывных функций.

23. Основные свойства интеграла. Установим ряд важных свойств определённого интеграла. Большая часть этих свойств присуща интегралам от любых интегрируемых функций, но мы будем формулировать их для функций непрерывных.

Теорема 1. *Если $f(x)$ и $g(x)$ — две непрерывные функции, заданные на промежутке $[a, b]$, то*

$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

т. е. интеграл суммы равен сумме интегралов слагаемых.

В самом деле, составляя интегральную сумму для функции $f(x)+g(x)$, очевидно, будем иметь

$$\sum_{k=0}^{n-1}[f(\xi_k)+g(\xi_k)](x_{k+1}-x_k)=$$

$$=\sum_{k=0}^{n-1}f(\xi_k)(x_{k+1}-x_k)+\sum_{k=0}^{n-1}g(\xi_k)(x_{k+1}-x_k),$$

после чего остаётся перейти к пределу при $\lambda\to 0$.

Аналогично доказывается

Теорема 2. *Если $f(x)$ — непрерывная функция, а c — постоянное число, то*

$$\int_a^b cf(x)\,dx=c\int_a^b f(x)\,dx,$$

т. е. постоянный множитель можно выносить за знак интеграла.

Теорема 3. *Пусть $f(x)$ непрерывна на промежутке $[a, b]$. Если этот промежуток точкой c разложен на части $[a, c]$ и $[c, b]$, то интеграл по всему промежутку оказывается равным сумме интегралов по его частям, т. е.*

$$\int_a^b f(x)\,dx=\int_a^c f(x)\,dx+\int_c^b f(x)\,dx \qquad (a<c<b).$$

В самом деле, будем при раздроблении промежутка $[a, b]$ на части включать точку c в число точек деления. Если $c=x_m$, то

$$\sum_{k=0}^{n-1}f(\xi_k)(x_{k+1}-x_k)=\sum_{k=0}^{m-1}f(\xi_k)(x_{k+1}-x_k)+\sum_{k=m}^{n-1}f(\xi_k)(x_{k+1}-x_k).$$

Каждая из написанных здесь трёх сумм является интегральной суммой соответственно для промежутков $[a, b]$, $[a, c]$ и $[c, b]$. Остаётся перейти к пределу при $\lambda\to 0$.

Доказанную теорему можно высказать в более общей форме. Для этого нам понадобится расширить смысл символа интеграла.

Если $f(x)$ — любая функция, определённая в точке a, то по определению полагаем

$$\int_a^a f(x)\,dx=0. \tag{11}$$

Таким образом, *интеграл с совпадающими пределами равен нулю.*

Пусть функция $f(x)$ интегрируема на промежутке $[a, b]$. Тогда мы по определению полагаем

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx. \tag{12}$$

Таким образом, *при перестановке пределов интегрирования определённый интеграл меняет знак.*

Теперь мы можем привести упомянутую более общую форму теоремы 3:

Теорема 4. *Пусть функция $f(x)$ непрерывна в промежутке $[A, B]$. Если a, b, c суть точки этого промежутка, то*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \tag{13}$$

В самом деле, если из точек a, b и c две (а тем более три) совпадают, то равенство (13) очевидно. Пусть же все эти точки различны. Если $a < c < b$, то дело сводится к теореме 3. Прочие случаи взаимного расположения точек a, b, c тоже легко свести к той же теореме. Пусть, например, $c < b < a$. Тогда

$$\int_c^a f(x)\,dx = \int_c^b f(x)\,dx + \int_b^a f(x)\,dx,$$

откуда

$$-\int_b^a f(x)\,dx = -\int_c^a f(x)\,dx + \int_c^b f(x)\,dx,$$

и остаётся дважды применить формулу (12).

Свойство интеграла, выражаемое теоремами 3 и 4, называется *аддитивностью* его, как функции промежутка интегрирования.

Теорема 5. *Если $f(x)$ — непрерывная функция, заданная на промежутке $[a, b]$, то существует такая точка $\xi \in [a, b]$, что*

$$\int_a^b f(x)\,dx = f(\xi)(b - a). \tag{14}$$

В самом деле, пусть M и m суть наибольшее и наименьшее значения $f(x)$ на промежутке $[a, b]$. Составим для $f(x)$ какую-нибудь интегральную сумму

$$\sigma = \sum_{k=0}^{n-1} f(\xi_k)(x_{k+1} - x_k).$$

Так как при всех k будет $m \leqslant f(\xi_k) \leqslant M$, а $x_{k+1} > x_k$, то $m(x_{k+1} - x_k) \leqslant f(\xi_k)(x_{k+1} - x_k) \leqslant M(x_{k+1} - x_k)$. Складывая такие неравенства и замечая, что

$$\sum_{k=0}^{n-1} (x_{k+1} - x_k) = b - a,$$

получим:

$$m(b-a) \leqslant \sigma \leqslant M(b-a).$$

Переходя в этом неравенстве к пределу при $\lambda \to 0$, приходим после деления на $b - a$ к новому неравенству

$$m \leqslant \frac{1}{b-a} \int_a^b f(x)\,dx \leqslant M.$$

Таким образом, частное

$$h = \frac{1}{b-a} \int_a^b f(x)\,dx$$

есть число, лежащее между наибольшим и наименьшим значениями непрерывной функции. Как известно, тогда и само это число должно являться одним из значений той же функции. Поэтому в $[a, b]$ обязательно существует такая точка ξ, что $h = f(\xi)$, а это равносильно равенству (14).

Заметим, что равенство (14) справедливо не только при $a < b$, но и при $a = b$ (тогда обе части этого равенства суть нули), а также и при $a > b$ (этот случай приводится к рассмотренному изменением знаков). В первом из этих случаев будет $\xi = a$, а во втором $a \geqslant \xi \geqslant b$.

Теорему 5 обычно называют *теоремой о среднем значении*. Из неё вытекает ряд свойств интеграла, выражающихся неравенствами.

Теорема 6. *Если $f(x)$ — неотрицательная непрерывная функция и нижний предел интеграла не больше верхнего* [1]), *то и сам интеграл будет числом неотрицательным*

$$\int_a^b f(x)\,dx \geqslant 0.$$

Действительно, в этом случае оба сомножителя правой части формулы (14) неотрицательны.

---

[1]) Если в интеграле $\int_a^b f(x)\,dx$ будет $a \leqslant b$, то мы будем говорить, что порядок пределов интегрирования — нормальный.

Последний результат можно несколько уточнить.

Теорема 7. *Если $a<b$, а $f(x)$ — непрерывная неотрицательная функция, которая хотя бы в одной точке $[a, b]$ отлична от нуля, то*

$$\int_a^b f(x)\,dx>0.$$

В самом деле, пусть $x_0\,(a<x_0<b)$ — такая точка, что $f(x_0)>0$. Возьмём столь малое $\delta>0$, чтобы при $|x-x_0|<\delta$ было $f(x)>0$, что, очевидно, возможно, благодаря непрерывности нашей функции. Не ограничивая общности, можно принять, что $a\leqslant x_0-\delta$, $x_0+\delta\leqslant b$. Тогда

$$\int_a^b f(x)\,dx=\int_a^{x_0-\delta} f(x)\,dx+\int_{x_0-\delta}^{x_0+\delta} f(x)\,dx+\int_{x_0+\delta}^b f(x)\,dx.$$

Первый и третий интегралы правой части по предыдущей теореме неотрицательны, а второй интеграл по теореме о среднем представим в форме

$$\int_{x_0-\delta}^{x_0+\delta} f(x)\,dx=f(\xi)\cdot 2\delta \qquad (x_0-\delta\leqslant\xi\leqslant x_0+\delta),$$

и потому строго положителен.

Теорему 7 можно, очевидно, формулировать и так:

Теорема 8. *Пусть $f(x)$ — неотрицательная непрерывная функция, заданная в $[a, b]$, причём $a<b$. Если*

$$\int_a^b f(x)\,dx=0,$$

то $f(x)$ всюду на $[a, b]$ равна нулю.

В обеих теоремах 7 и 8 (в отличие от теоремы 6) нельзя отбросить условия непрерывности подинтегральной функции. Например, функция, которая в конечном числе точек $[a, b]$ равна единице, а в остальных точках этого промежутка равна нулю, будет неотрицательной и нетождественной нулю, а интеграл от неё (как было показано в п° 22) равен нулю.

Теорема 9. *Если $a\leqslant b$, а $f(x)$ и $g(x)$ — две непрерывные функции, которые на $[a, b]$ удовлетворяют условию $f(x)\leqslant g(x)$, то*

$$\int_a^b f(x)\,dx\leqslant\int_a^b g(x)\,dx, \tag{15}$$

т. е. при нормальном порядке пределов интегрирования неравенство можно интегрировать почленно.

Действительно,

$$\int_a^b g(x)\,dx - \int_a^b f(x)\,dx = \int_a^b [g(x) - f(x)]\,dx \geqslant 0.$$

Если бы мы допустили, что $a < b$ и что хоть в одной точке оказывается $f(x) < g(x)$, то смогли бы в (15) исключить знак равенства.

Теорема 10. *Если $a \leqslant b$ и $f(x)$ непрерывна на $[a, b]$, то*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx, \tag{16}$$

т. е. при нормальном порядке пределов интегрирования абсолютная величина интеграла не превосходит интеграла от абсолютной величины подинтегральной функции.

В самом деле, интегрируя неравенство

$$-|f(x)| \leqslant f(x) \leqslant |f(x)|,$$

находим:

$$-\int_a^b |f(x)|\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b |f(x)|\,dx,$$

а это равносильно неравенству (16).

Из теоремы о среднем значении немедленно вытекает также следующая важная оценка интеграла:

Теорема 11. *Если непрерывная функция $f(x)$ для всех x между a и b удовлетворяет неравенству $|f(x)| \leqslant K$, то*

$$\left|\int_a^b f(x)\,dx\right| \leqslant K|b - a|.$$

Отметим, наконец, что поскольку определённый интеграл есть постоянное число, вполне определяемое пределами интегрирования и подинтегральной функцией, то *обозначение переменной интегрирования никакого значения иметь не может,* так что символы

$$\int_a^b f(x)\,dx, \quad \int_a^b f(z)\,dz, \quad \int_a^b f(t)\,dt, \quad \int_a^b f(u)\,du$$

обозначают одно и то же число. Это небесполезно сопоставить с тем, что у интегралов неопределённых дело обстоит не так. Например,

$$\int x^2\,dx = \frac{x^3}{3} + C, \quad \int z^2\,dz = \frac{z^3}{3} + C.$$

Поэтому, сводя интеграл

$$\int \sin^2 x \cos x \, dx$$

с помощью подстановки $\sin x = z$ к интегралу

$$\int z^2 \, dz,$$

мы должны иметь в виду, что последний интеграл равен не $\frac{x^3}{3} + C$, а именно, $\frac{z^3}{3} + C$ с тем, чтобы в этом выражении заменить z на $\sin x$.

24. Интеграл, как функция верхнего предела. До сих пор мы рассматривали свойства определённого интеграла, считая пределы интегрирования постоянными. Теперь же мы рассмотрим вопрос о том, как влияет изменение этих пределов на величину интеграла.

Пусть $f(x)$ — непрерывная функция, заданная на промежутке $[a, b]$. Тогда она будет непрерывной и на всяком частичном промежутке $[a, x]$, и мы можем рассмотреть интеграл

$$\int_a^x f(t) \, dt \qquad (a \leqslant x \leqslant b),$$

являющийся функцией аргумента x (как указывалось в конце предыдущего п°, обозначение переменной интегрирования не существенно. Чтобы не путать эту переменную с пределом интегрирования, мы обозначаем её через t).

Имеет место замечательная теорема, которую следует считать основной теоремой математического анализа:

Теорема. *Производная определённого интеграла от непрерывной функции, рассматриваемого как функция его верхнего предела, существует и равна значению подинтегральной функции в точке дифференцирования.*

В виде формулы высказанное утверждение выглядит так:

$$\left(\int_a^x f(t) \, dt \right)' = f(x).$$

Если же положить

$$\int_a^x f(t) \, dt = \Phi(x),$$

то формулированную теорему можно будет записать равенством

$$\Phi'(x) = f(x). \tag{17}$$

Приведём сначала не строгое, но очень наглядное геометрическое рассуждение, выясняющее суть теоремы.

Предполагая функцию $f(t)$ непрерывной и положительной на $[a, b]$, мы сможем изобразить функцию $\Phi(x)$ в виде площади криволинейной трапеции, ограниченной линиями $y=0$, $t=a$, $t=x$, $y=f(t)$ (рис. 25). Придавая аргументу x приращение Δx (пусть для простоты $\Delta x>0$), мы изобразим приращение $\Delta\Phi$ функции $\Phi(x)$ в виде площади узкой полоски, заштрихованной на чертеже.

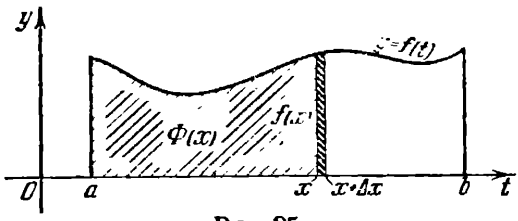


Рис. 25.

Приняв эту полоску за прямоугольник с основанием Δx и высотой $f(x)$ (где x — точка дифференцирования), мы получаем приближённое равенство

$$\Delta\Phi \approx f(x)\,\Delta x,$$

которое можно записать и так:

$$\frac{\Delta\Phi}{\Delta x} \approx f(x).$$

Так как точность этого равенства тем выше, чем меньше Δx, то

$$f(x) = \lim_{\Delta x \to 0} \frac{\Delta\Phi}{\Delta x},$$

а это равносильно формуле (17).

Переходя к точному доказательству основной теоремы, рассмотрим две точки x и $x+\Delta x$ из промежутка $[a, b]$, на котором задана непрерывная функция $f(t)$. Тогда

$$\Phi(x) = \int_a^x f(t)\,dt,$$

$$\Phi(x+\Delta x) = \int_a^{x+\Delta x} f(t)\,dt = \int_a^x f(t)\,dt + \int_x^{x+\Delta x} f(t)\,dt.$$

Отсюда, применяя теорему о среднем значении, находим

$$\Phi(x+\Delta x) - \Phi(x) = \int_x^{x+\Delta x} f(t)\,dt = f(\xi)\,\Delta x,$$

причём ξ лежит между x и $x+\Delta x$. В таком случае

$$\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x} = f(\xi).$$

Если $\Delta x \to 0$, то точка ξ стремится к точке x. В силу непрерывности функции $f(t)$ отсюда следует, что $f(\xi)$ стремится к $f(x)$. Поэтому

$$\lim_{\Delta x \to 0} \frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x} = f(x).$$

Теорема доказана. Из неё вытекает, в частности, что $\Phi(x)$ есть функция непрерывная.

Примерами, иллюстрирующими доказанную теорему, могут служить следующие равенства:

$$\left(\int_2^x e^{t^2}\,dt\right)' = e^{x^2}, \quad \left(\int_0^x \sqrt{\cos z}\,dz\right)' = \sqrt{\cos x}, \quad \left(\int_1^z \frac{dx}{x}\right)' = \frac{1}{z}.$$

В п° 17 мы сообщили без доказательства, что у всякой непрерывной на каком-нибудь промежутке функции существует на этом промежутке первообразная. Теперь это утверждение становится очевидным. Действительно, если $f(x)$ непрерывна на промежутке $[a, b]$, то функция

$$\Phi(x) = \int_a^x f(t)\,dt$$

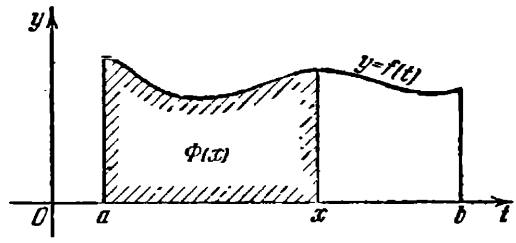


Рис. 26.

служит для неё первообразной. Более того, мы имеем и геометрическое изображение этой первообразной на графике самой функции $y = f(t)$ (рис. 26).

25. Вычисление определённого интеграла с помощью неопределённого. Пусть функция $f(x)$ непрерывна на промежутке $[a, b]$. Займёмся вопросом о вычислении интеграла

$$I = \int_a^b f(x)\,dx.$$

Для решения этого вопроса прежде всего перепишем интеграл I в форме

$$I = \int_a^b f(t)\,dt,$$

введя новое обозначение для переменной интегрирования. Теперь вместо интеграла I мы рассмотрим функцию

$$\Phi(x) = \int_a^x f(t)\,dt.$$

Легко видеть, что $I=\Phi(b)$, так что, найдя функцию $\Phi(x)$, мы и подавно сможем вычислить интересующий нас интеграл I. На первый взгляд, поставив вопрос о нахождении функции $\Phi(x)$ вместо индивидуального постоянного числа I, мы усложнили стоящую перед нами задачу. На самом деле, однако, такое разумное обобщение вопроса позволяет применить для его решения такие методы, которые нельзя было бы использовать при первоначальной частной постановке задачи. Следует заметить, что в науке неоднократно встречается подобное положение — более общая постановка вопроса создаёт больше путей подхода к его решению.

В частности, в нашем случае получается возможность опереться на теорему предыдущего n°, согласно которой

$$\Phi'(x)=f(x).$$

Таким образом, функция $\Phi(x)$ оказывается одной из первообразных функций для подинтегральной функции $f(x)$. Допустим, что нам известна какая-нибудь из первообразных для $f(x)$, например $F(x)$. Тогда полное семейство первообразных для $f(x)$ даётся формулой

$$\int f(x)\,dx=F(x)+C, \tag{18}$$

и так как $\Phi(x)$ содержится в этом семействе, то

$$\Phi(x)=F(x)+C_0, \tag{19}$$

где C_0 есть именно то значение постоянной C, при котором из полного семейства первообразных получается $\Phi(x)$.

Если мы вспомним, что при совпадении пределов интегрирования интеграл обращается в нуль, то сразу увидим, что

$$\Phi(a)=\int_a^a f(t)\,dt=0.$$

Отсюда и из (19) вытекает, что

$$C_0=-F(a),$$

и потому

$$\Phi(x)=F(x)-F(a).$$

В частности, при $x=b$ находим:

$$\int_a^b f(x)\,dx=F(b)-F(a).$$

Эта формула (называемая *формулой Ньютона-Лейбница*) сводит вопрос о вычислении определённого интеграла любой непрерывной

функции к нахождению для неё первообразной функции. По существу этим перекинут мост между двумя частями математического анализа — дифференциальным исчислением (к которому, собственно, надо отнести и понятие первообразной функции) и интегральным исчислением, которое изучает в основном пределы интегральных сумм. К концу XVII в. оба эти исчисления были разработаны уже весьма обстоятельно, но то, что они связаны между собой, ещё не было выяснено. Заслугой Ньютона и Лейбница является именно установление факта этой связи. Мы видим, что в основе её лежит предложение, составляющее содержание теоремы n° 24, почему мы и назвали эту теорему основной теоремой математического анализа.

Ввиду чрезвычайной важности установленного результата придадим ему форму следующего правила:

Правило. *Для вычисления определённого интеграла от непрерывной функции надо найти для неё первообразную функцию и составить разность значений этой последней функции при верхнем и нижнем пределах интегрирования.*

При выводе этого правила и выражающей его формулы (20) мы считали, что $a<b$. Однако это не существенно. Действительно, при $a=b$ формула (20) очевидна, ибо обе её части равны нулю. Случай же $a>b$ приводится к случаю $a<b$ переменой знака обеих частей формулы (20).

Формулу (20) можно переписать иначе, если ввести очень удобное обозначение разности $F(b)-F(a)$ символом

$$[F(x)]_a^b.$$

При этом обозначении формула (20) принимает вид

$$\int_a^b f(x)\,dx=[F(x)]_a^b. \tag{21}$$

Если заметить, что в качестве $F(x)$ может быть использована любая первообразная $F(x)+C$, то формулу Ньютона-Лейбница можно будет записать и так:

$$\int_a^b f(x)\,dx=\left[\int f(x)\,dx\right]_a^b.$$

Это, пожалуй, наиболее выразительная её запись.

Приведём несколько примеров:

1) $$\int_{\frac{\pi}{6}}^{\frac{\pi}{2}}\cos x\,dx=\left[\sin x\right]_{\frac{\pi}{6}}^{\frac{\pi}{2}}=\sin\frac{\pi}{2}-\sin\frac{\pi}{6}=1-\frac{1}{2}=\frac{1}{2},$$

$$2)\ \int_1^4 x^2\,dx = \left[\frac{x^3}{3}\right]_1^4 = \frac{4^3}{3} - \frac{1^3}{3} = 21,$$

$$3)\ \int_0^1 \frac{dx}{1+x^2} = [\operatorname{arctg} x]_0^1 = \operatorname{arctg} 1 - \operatorname{arctg} 0 = \frac{\pi}{4}.$$

Само собой разумеется, что все приёмы, позволяющие вычислять неопределённые интегралы, применимы и тогда, когда речь идёт о вычислении определённого интеграла. В частности, для вычисления определённого интеграла можно применять замену переменной и интегрирование по частям. Однако в применении к задаче вычисления определённого интеграла указанные два приёма могут быть несколько специализированы.

Теорема 1. *Пусть $f(z)$ — непрерывная функция, заданная на промежутке $[p, q]$, а $\varphi(x)$ — непрерывная функция, заданная на промежутке $[a, b]$, имеющая там непрерывную же производную $\varphi'(x)$ и удовлетворяющая неравенству $p \leqslant \varphi(x) \leqslant q$.*

В таком случае

$$\int_a^b f[\varphi(x)]\,\varphi'(x)\,dx = \int_{\varphi(a)}^{\varphi(b)} f(z)\,dz. \tag{22}$$

Формула (22) выражает собой правило замены переменной в определённом интеграле. Оно напоминает правило замены переменной в интеграле неопределённом, но отличается от него тем, что здесь отпадает надобность в возвращении к старой переменной, ибо формула (22) представляет собой равенство двух постоянных чисел. Заметим ещё, что эта формула заменяет собой для случая определённых интегралов оба вида правила подстановки в интегралах неопределённых; только, применяя её на практике, иной раз приходится читать её слева направо, а иногда — справа налево.

Переходя к доказательству теоремы, обозначим интегралы, входящие в левую и правую части формулы (22), соответственно через $I_{\text{лев}}$ и $I_{\text{прав}}$.

Пусть $F(z)$ — функция первообразная для $f(z)$. Тогда по формуле Ньютона-Лейбница

$$I_{\text{прав}} = F[\varphi(b)] - F[\varphi(a)]. \tag{23}$$

Что же касается $I_{\text{лев}}$, то

$$I_{\text{лев}} = \left[\int f[\varphi(x)]\,\varphi'(x)\,dx\right]_a^b.$$

Но согласно теореме из п° 18 будет

$$\int f[\varphi(x)]\,\varphi'(x)\,dx = F[\varphi(x)] + C.$$

Значит,

$$I_{\text{лев}} = F[\varphi(b)] - F[\varphi(a)].$$

Отсюда и из (23) следует, что $I_{\text{лев}} = I_{\text{прав}}$.

Пример 1.

$$\int_1^e \frac{dx}{x(1+\ln^2 x)} = \int_0^1 \frac{dz}{1+z^2} = \frac{\pi}{4}.$$

Здесь применена подстановка $\ln x = z$ (причём формула (22) прочитывалась слева направо).

Пример 2.

$$\int_0^R \sqrt{R^2 - x^2}\, dx = \int_0^{\frac{\pi}{2}} R^2 \cos^2 t\, dt.$$

Здесь применена подстановка $x = R\sin t$, и формула (22) прочитывается справа налево. Дальнейшее вычисление просто:

$$\int_0^{\frac{\pi}{2}} \cos^2 t\, dt = \int_0^{\frac{\pi}{2}} \frac{1+\cos 2t}{2}\, dt = \left[\frac{t}{2} + \frac{\sin 2t}{4}\right]_0^{\frac{\pi}{2}} = \frac{\pi}{4},$$

и окончательно

$$\int_0^R \sqrt{R^2 - x^2}\, dx = \frac{1}{4}\pi R^2.$$

Теорема 2. *Если $u(x)$ и $v(x)$ — две функции, заданные на промежутке $[a, b]$ и имеющие там непрерывные производные, то*

$$\int_a^b u\, dv = [uv]_a^b - \int_a^b v\, du. \tag{24}$$

Формула (24) есть формула интегрирования по частям для определённых интегралов.

Доказательство очень просто. Именно,

$$\int_a^b u\, dv = \left[\int u\, dv\right]_a^b.$$

Так как по формуле интегрирования по частям будет

$$\int u\,dv = uv - \int v\,du,$$

то

$$\int_a^b u\,dv = \left[uv - \int v\,du\right]_a^b,$$

откуда и следует (24).

Пример 3. Пусть

$$I = \int_0^{\pi} x\cos x\,dx.$$

Тогда

$$I = \int_0^{\pi} x\,d\sin x = [x\sin x]_0^{\pi} - \int_0^{\pi}\sin x\,dx = [\cos x]_0^{\pi} = -2.$$

26. Формула Валлиса. С помощью результатов предыдущих пп° можно вывести интересное выражение для числа π. Рассмотрим интеграл

$$U_n = \int_0^{\frac{\pi}{2}} \cos^n x\,dx,$$

где n — целое и неотрицательное число. Легко видеть, что

$$U_0 = \frac{\pi}{2}, \quad U_1 = 1.$$

Предположим теперь, что $n > 1$. Тогда, интегрируя по частям, мы находим:

$$U_n = \int_0^{\frac{\pi}{2}} \cos^{n-1} x\,d\sin x = \left[\sin x\cos^{n-1}x\right]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} \sin x\,d(\cos^{n-1}x).$$

Так как первый член правой части равен нулю, то

$$U_n = (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}x\sin^2 x\,dx = (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}x\,(1-\cos^2 x)\,dx.$$

Отсюда

$$U_n = (n-1)\,(U_{n-2} - U_n)$$

и, стало быть,

$$U_n = \frac{n-1}{n}\,U_{n-2}. \tag{25}$$

Допустим, что n — число нечётное. Если $n-2>1$, то мы можем снова написать формулу (25), изменив только n на $n-2$, что даёт

$$U_{n-2}=\frac{n-3}{n-2}U_{n-4}.$$

Так можно продолжать до тех пор, пока мы не придём к равенству

$$U_3=\frac{2}{3}U_1.$$

Выражая U_n через U_{n-2}, затем U_{n-2} через U_{n-4} и т. д., получим:

$$U_n=\frac{2\cdot 3\cdot 4\cdot\ldots\cdot(n-1)}{3\cdot 4\cdot 5\cdot\ldots\cdot n}U_1.$$

Вспоминая, что $U_1=1$, и обозначая произведение всех натуральных чисел, не превосходящих m и имеющих с m одинаковую чётность, через $m!!$, находим, что

$$U_n=\frac{(n-1)!!}{n!!}. \tag{26}$$

Аналогично для чётного n мы получим:

$$U_n=\frac{(n-1)!!}{n!!}\frac{\pi}{2}. \tag{27}$$

Установив эти равенства, заметим, что при $0\leqslant x\leqslant\frac{\pi}{2}$ будет

$$\cos^{2n+2}x\leqslant\cos^{2n+1}x\leqslant\cos^{2n}x,$$

откуда, интегрируя и опираясь на теорему 9 из п° 23, находим:

$$U_{2n+2}<U_{2n+1}<U_{2n}.$$

С помощью равенств (26) и (27) последнему неравенству можно дать вид

$$\frac{(2n+1)!!}{(2n+2)!!}\frac{\pi}{2}<\frac{(2n)!!}{(2n+1)!!}<\frac{(2n-1)!!}{(2n)!!}\frac{\pi}{2}.$$

Отсюда

$$\frac{2n+1}{2n+2}\cdot\frac{\pi}{2}<\left[\frac{(2n)!!}{(2n-1)!!}\right]^2\cdot\frac{1}{2n+1}<\frac{\pi}{2}. \tag{28}$$

Так как

$$\lim_{n\to\infty}\frac{2n+1}{2n+2}=1,$$

то из (28) следует, что

$$\frac{\pi}{2}=\lim_{n\to\infty}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2\cdot\frac{1}{2n+1}.$$

Отсюда в свою очередь вытекает, что

$$\pi=\lim_{n\to\infty}\frac{1}{n}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2.$$

Эта формула носит название *формулы Валлиса*. Она даёт довольно простое выражение числа π через натуральные числа. Теоретически этот

результат интересен [1]). Что же касается ценности этой формулы как средства фактического вычисления π, то она невелика. Именно, чтобы получить удовлетворительную точность, надо взять n довольно большим, а тогда выражение $\frac{(2n)!!}{(2n-1)!!}$ оказывается весьма громоздким.

27. Приближённое вычисление определённых интегралов. Формула Ньютона-Лейбница позволяет нам вычислять определённые интегралы от таких функций, первообразные которых выражаются конечным числом элементарных функций. В тех же случаях, когда упомянутая первообразная не выражается через элементарные функции (или когда нахождение её связано с чрезмерно громоздкими вычислениями), приходится искать иные способы вычисления определённого интеграла.

В исключительных случаях удаётся найти определённый интеграл каким-нибудь искусственным приёмом.

Например, сделав в интеграле [2])

$$I=\int_0^{\pi}\frac{x\sin x}{1+\cos^2 x}\,dx$$

подстановку $x=\pi-z$, приведём его к виду

$$I=\int_0^{\pi}\frac{(\pi-z)\sin z}{1+\cos^2 z}\,dz,$$

откуда

$$I=\pi\int_0^{\pi}\frac{\sin z\,dz}{1+\cos^2 z}-I.$$

Значит,

$$2I=\pi\int_0^{\pi}\frac{\sin z\,dz}{1+\cos^2 z}.$$

Положив здесь $\cos z=u$, получаем:

$$2I=\pi\int_{-1}^{1}\frac{du}{1+u^2}=\pi\,[\operatorname{arctg} u]_{-1}^{+1}=\frac{\pi^2}{2}$$

и окончательно

$$\int_0^{\pi}\frac{x\sin x}{1+\cos^2 x}\,dx=\frac{\pi^2}{4}.$$

Однако ясно, что такие случаи не типичны. Вообще же говоря, определённый интеграл от такой функции, первообразная которой

---

[1]) Ниже будут даны и другие выражения π через натуральные числа.

[2]) Отметим, что неопределённый интеграл $\int\frac{x\sin x}{1+\cos^2 x}\,dx$ не выражается через элементарные функции.

не выражается через элементарные функции, приходится находить с помощью какой-либо приближённой формулы.

Мы остановимся только на одной из таких формул, которая называется «формулой средних прямоугольников». Разнообразные другие формулы приближённого интегрирования основаны по существу на тех же принципах, но потребовали бы бо́льших усилий для установления оценки доставляемой ими точности.

Как мы установили выше, для интеграла от непрерывной функции справедлива формула

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

Формула эта абсолютно то́чна, но не даёт способа вычислять интеграл, ибо точка ξ нам неизвестна. Заменим в этой формуле неизвестную точку ξ на середину промежутка $[a, b]$. Это приведёт нас уже не к точной, а только к приближённой формуле

$$\int_a^b f(x)\,dx = f\left(\frac{a+b}{2}\right)(b-a), \tag{29}$$

которая называется «малой формулой средних прямоугольников». Легко понять, что с геометрической точки зрения замена интеграла

$$\int_a^b f(x)\,dx$$

величиной $f\left(\frac{a+b}{2}\right)(b-a)$ означает замену криволинейной трапеции, ограниченной линиями $y=0$, $y=f(x)$, $x=a$, $x=b$, прямоугольником с основанием $[a, b]$ и высотой $f\left(\frac{a+b}{2}\right)$ (рис. 27).

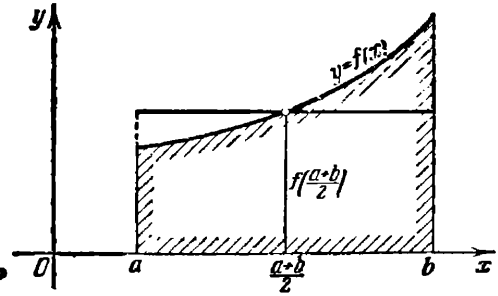


Рис. 27.

Предположим, что функция $f(x)$ имеет на $[a, b]$ непрерывные первую и вторую производные $f'(x)$ и $f''(x)$, и установим, какова при этом оказывается ошибка формулы (29).

Разлагая $f(x)$ по формуле Тейлора по степеням разности $x-\frac{a+b}{2}$, находим

$$f(x) = f\left(\frac{a+b}{2}\right) + f'\left(\frac{a+b}{2}\right)\left(x-\frac{a+b}{2}\right) + \frac{f''(\bar{x})}{2}\left(x-\frac{a+b}{2}\right)^2,$$

где $\bar{x}$ — некоторая неизвестная нам точка, лежащая между $\frac{a+b}{2}$ и x.

Интегрируя это равенство и замечая, что

$$\int_a^b \left(x - \frac{a+b}{2}\right) dx = 0,$$

получаем:

$$\int_a^b f(x)\,dx = f\left(\frac{a+b}{2}\right)(b-a) + \frac{1}{2}\int_a^b f''(\bar{x})\left(x - \frac{a+b}{2}\right)^2 dx. \quad (30)$$

Обозначим через m и M наименьшее и наибольшее значения непрерывной функции $f''(x)$. Тогда

$$m \leqslant f''(\bar{x}) \leqslant M.$$

Умножая это неравенство на (неотрицательную!) величину $\left(x - \frac{a+b}{2}\right)^2$ и интегрируя, находим:

$$m\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \leqslant \int_a^b f''(\bar{x})\left(x - \frac{a+b}{2}\right)^2 dx \leqslant M\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx.$$

Интеграл

$$\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx$$

с помощью подстановки $x = \frac{a+b}{2} + \frac{b-a}{2}t$ приводится к интегралу

$$\frac{(b-a)^3}{8}\int_{-1}^{1} t^2\,dt = \frac{(b-a)^3}{12}.$$

Поэтому

$$m\frac{(b-a)^3}{12} \leqslant \int_a^b f''(\bar{x})\left(x - \frac{a+b}{2}\right)^2 dx \leqslant M\frac{(b-a)^3}{12}.$$

Отсюда следует, что число

$$h = \frac{12}{(b-a)^3}\int_a^b f''(\bar{x})\left(x - \frac{a+b}{2}\right)^2 dx$$

лежит между m и M. Так как $f''(x)$ есть функция непрерывная, то она принимает значение h при каком-то значении аргумента $x = \eta$. Значит,

$$\int_a^b f''(\bar{x})\left(x - \frac{a+b}{2}\right)^2 dx = \frac{(b-a)^3}{12}f''(\eta).$$

Подставляя эту величину в (30), мы получаем «малую формулу средних прямоугольников с остаточным членом»:

$$\int_a^b f(x)\,dx = f\left(\frac{a+b}{2}\right)(b-a) + \frac{(b-a)^3}{24} f''(\eta). \qquad (31)$$

Установив этот результат, представим себе, что промежуток интегрирования $[a, b]$ разбит точками

$$x_0 = a, \quad x_1 = a + \frac{b-a}{n}, \quad x_2 = a + 2\frac{b-a}{n}, \ \ldots, \ x_n = b$$

на n равных частей $[x_k, x_{k+1}]$. Тогда

$$\int_a^b f(x)\,dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x)\,dx. \qquad (32)$$

Применяя формулу (31) к интегралам по частичным промежуткам, находим:

$$\int_{x_k}^{x_{k+1}} f(x)\,dx = f\left(\frac{x_k + x_{k+1}}{2}\right)(x_{k+1} - x_k) + \frac{(x_{k+1} - x_k)^3}{24} f''(\eta_k), \qquad (33)$$

где $x_k \leqslant \eta_k \leqslant x_{k+1}$. Замечая, что $x_{k+1} - x_k = \frac{b-a}{n}$ и обозначая точку $\frac{x_k + x_{k+1}}{2}$ через $x_{k+1/2}$, находим из (32) и (33):

$$\int_a^b f(x)\,dx = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_{k+1/2}) + \frac{(b-a)^3}{24n^3} \sum_{k=0}^{n-1} f''(\eta_k). \qquad (34)$$

Но так как значения $f''(\eta_k)$ лежат между m и M, то

$$m \leqslant \frac{1}{n} \sum_{k=0}^{n-1} f''(\eta_k) \leqslant M.$$

Поэтому существует такая точка ξ, что

$$f''(\xi) = \frac{1}{n} \sum_{k=0}^{n-1} f''(\eta_k)$$

и равенство (6) принимает вид

$$\int_a^b f(x)\,dx = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_{k+1/2}) + \frac{(b-a)^3}{24n^2} f''(\xi). \qquad (35)$$

Это — «большая формула средних прямоугольников с остаточным членом». Отбрасывая последнее слагаемое, мы получим приближённую формулу средних прямоугольников:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n}\left[f(x_{1/2})+f(x_{3/2})+\ldots+f(x_{n-1/2})\right]. \quad (36)$$

Из изложенного ясно, что абсолютная величина ошибки этой формулы не больше, чем

$$K\frac{(b-a)^3}{24n^2}, \quad (37)$$

где $K=\max|f''(x)|$.

Величина (37) с возрастанием n стремится к нулю и потому

$$\int_a^b f(x)\,dx=\lim_{n\to\infty}\frac{b-a}{n}\left[f(x_{1/2})+f(x_{3/2})+\ldots+f(x_{n-1/2})\right].$$

Эта последняя формула, впрочем, верна вообще для всех интегрируемых функций [1]), но лишь для функций с непрерывной второй производной мы можем оценить достигнутую (при заданном n) точность.

Рассмотрим несколько примеров применения выведенной формулы.

Пример 1. Вычислить интеграл

$$\int_0^1 x^2\,dx$$

по формуле (36), взяв $n=10$.

Здесь $x_k=\frac{k}{10}$. Значит,

$$x_{k+1/2}=\frac{x_k+x_{k+1}}{2}=\frac{2k+1}{20}.$$

---

[1]) Действительно, $\frac{b-a}{n}\left[f(x_{1/2})+\ldots+f(x_{n-1/2})\right]$ есть интегральная сумма, отвечающая дроблению $[a, b]$ точками

$$x_k=a+\frac{k}{n}(b-a)$$

и выбору точек

$$\xi_k=x_{k+1/2}=\frac{x_k+x_{k+1}}{2}.$$

Поэтому

$$\left.\begin{aligned} &x_{1/2}=0{,}05,\ x_{3/2}=0{,}15,\ x_{5/2}=0{,}25,\ x_{7/2}=0{,}35,\ x_{9/2}=0{,}45,\\ &x_{11/2}=0{,}55,\ x_{13/2}=0{,}65,\ x_{15/2}=0{,}75,\ x_{17/2}=0{,}85,\ x_{19/2}=0{,}95. \end{aligned}\right\} \quad (38)$$

Так как у нас $f(x)=x^2$, то

$$\begin{aligned} &f(x_{1/2})=0{,}0025,\quad f(x_{3/2})=0{,}0225,\quad f(x_{5/2})=0{,}0625,\\ &f(x_{7/2})=0{,}1225,\quad f(x_{9/2})=0{,}2025,\quad f(x_{11/2})=0{,}3025,\\ &f(x_{13/2})=0{,}4225,\quad f(x_{15/2})=0{,}5625,\quad f(x_{17/2})=0{,}7225,\\ &f(x_{19/2})=0{,}9025. \end{aligned}$$

Складывая эти значения, получаем:

$$\sum_{k=0}^{9} f(x_{k+1/2})=3{,}325$$

и формула (36) даёт

$$\int_0^1 x^2\,dx=0{,}3325. \quad (39)$$

В действительности, значение этого интеграла равно 0,33333. Значит, ошибка равенства (39) меньше, чем 0,001, а относительная его ошибка меньше, чем 0,003, т. е. $0{,}3\%$.

Пример 2. Вычислить

$$\int_0^1 \frac{dx}{1+x^2}$$

по формуле средних прямоугольников (36) при $n=10$.

Так как здесь рассматривается тот же промежуток [0, 1], что и в предыдущем примере, и взято то же значение n, то точками $x_{k+1/2}$ попрежнему являются точки (38).

Эти точки мы должны подставить в формулу

$$f(x)=\frac{1}{1+x^2}$$

и выразить результаты с помощью десятичных дробей. Чтобы установить, с каким количеством знаков надо писать эти дроби, оценим, какую ошибку мы делаем в нашем примере, применяя формулу (36) при $n=10$. У нас

$$f''(x)=\frac{2(3x^2-1)}{(1+x^2)^3}.$$

Ясно, что абсолютная величина этой дроби не больше[1]), чем 2. Значит, величина (37) в нашем примере равна $\frac{1}{1200}$. Вычисляя каждое значение $f(x_{k+1/2})$ по правилу дополнения с четырьмя знаками после запятой, мы сделаем в этом значении ошибку, меньшую чем 0,00005. Значит, сложив десять таких значений, мы ошибаемся меньше чем на 0,0005.

Эту сумму нам придётся умножить на $\frac{b-a}{n}=0{,}1$, а потому и ошибка её уменьшится в десять раз.

Таким образом, суммарная ошибка, происходящая и от погрешности формулы и от округления, будет всё же меньше, чем 0,001.

После этих замечаний можно перейти к вычислениям. Так как

$$f(x_{1/2})=0{,}9975,\quad f(x_{11/2})=0{,}7678,$$
$$f(x_{3/2})=0{,}9780,\quad f(x_{13/2})=0{,}7030,$$
$$f(x_{5/2})=0{,}9412,\quad f(x_{15/2})=0{,}6400,$$
$$f(x_{7/2})=0{,}8909,\quad f(x_{17/2})=0{,}5806,$$
$$f(x_{9/2})=0{,}8316,\quad f(x_{19/2})=0{,}5256,$$

то

$$\sum_{k=0}^{9} f(x_{k+1/2})=7{,}8562$$

и, стало быть,

$$\int_0^1 \frac{dx}{1+x^3}=0{,}7856 \qquad (\pm 0{,}001).$$

Замечая, что значение этого интеграла есть $\frac{\pi}{4}$, получаем:

$$\pi=3{,}1424\ (\pm 0{,}004).$$

Таким образом, вполне строго доказаны неравенства

$$3{,}138<\pi<3{,}147.$$

Как известно, на самом деле $\pi=3{,}14159$.

Рассмотренный пример имеет важное принципиальное значение, так как здесь мы видим практически применимый способ вычисления числа π (а ведь это одна из важнейших постоянных математики) с любой степенью точности (в п° 42 будут изложены другие, более быстрые способы вычисления π).

---

[1]) Действительно, $\left|\frac{3x^2-1}{(1+x^3)^2}\right|\leqslant\frac{3x^2+1}{1+3x^3+3x^6+x^6}\leqslant 1.$

Пример 3. Вычислить интеграл

$$\int_0^1 \frac{dx}{1+x}$$

по формуле (36), взяв $n=10$.

Здесь мы должны те же точки (38) подставлять в функцию

$$f(x)=\frac{1}{1+x}.$$

Результаты таковы:

$$\begin{array}{ll|l} f(x_{1/2})=0{,}9524, & f(x_{11/2})=0{,}6452, & \\ f(x_{3/2})=0{,}8696, & f(x_{13/2})=0{,}6061, & \sum_{k=0}^{9} f(x_{k+1/2})=6{,}9284, \\ f(x_{5/2})=0{,}8000, & f(x_{15/2})=0{,}5714, & \\ f(x_{7/2})=0{,}7407, & f(x_{17/2})=0{,}5405, & \frac{b-a}{n}\sum_{k=0}^{9} f(x_{k+1/2})=0{,}69284. \\ f(x_{9/2})=0{,}6897, & f(x_{19/2})=0{,}5128, & \end{array}$$

Ошибка, происходящая от округления, в каждом слагаемом меньше, чем 0,00005. Значит, сумма подсчитана с ошибкой, меньшей чем 0,0005, а величина $\frac{b-a}{n}\sum f(x_{k+1/2})$ получается с ошибкой, меньшей чем 0,00005. С другой стороны, остаточный член формулы прямоугольников оценивается числом (37). В нашем случае $f''(x)=\frac{2}{(1+x)^3}$, и потому можно взять $K=2$, откуда следует, что ошибка формулы меньше, чем

$$\frac{1}{1200}<0{,}00084.$$

Кроме того, в этом примере можно учесть ещё, что знак $f''(x)$ положительный. Таким образом, ошибка R формулы удовлетворяет неравенству

$$0<R<0{,}00084,$$

а суммарная ошибка лежит между пределами

$$-0{,}00005 \quad \text{и} \quad +0{,}00089.$$

Замечая, что интересующий нас интеграл равен $\ln 2$, находим:

$$0{,}69279<\ln 2<0{,}69373.$$

Значит, и подавно $\ln 2=0{,}693$ $(\pm 0{,}001)$.

Совершенно аналогично, применяя формулу (36) при достаточно большом n к интегралу

$$\int_0^{N-1} \frac{dx}{1+x} = \ln N,$$

мы сможем вычислить логарифм любого положительного числа N с любой нужной нам точностью (и оценить погрешность!). Таким образом, мы располагаем теперь эффективным способом составления логарифмических таблиц. Ниже (в п° 41) мы покажем ещё и другой подход к этому вопросу.

В заключение остановимся на вопросе о фактическом нахождении численных значений первообразной для непрерывной функции в том случае, когда эта первообразная не выражается элементарно. Легко видеть, что теперь мы имеем возможность находить эти значения. Действительно, пусть $f(x)$ есть непрерывная функция, заданная на промежутке $[a, b]$. Остановим своё внимание на той её первообразной, которая даётся формулой

$$\Phi(x) = \int_a^x f(t)\,dt \tag{40}$$

(всякая другая получается из неё прибавлением той или иной постоянной). Чтобы вычислить функцию $\Phi(x)$ для какого-либо x, надо вычислить интеграл, стоящий в равенстве (40), а это мы теперь уже умеем делать. Правда, чтобы получаемый результат был в какой-нибудь степени надёжен, надо уметь оценить его погрешность. Для произвольной непрерывной функции мы не располагаем способом оценки погрешности, но для функции с непрерывной второй производной такой способ у нас есть. В частности, поставленный вопрос решается до конца тогда, когда $f(x)$ есть элементарная функция (с непрерывной второй производной).

§ 6. Приложения интегрального исчисления

28. Вычисление площадей. Мы уже видели выше, что площадь F криволинейной трапеции, ограниченной линиями $y=0$, $x=a$, $x=b$ и $y=f(x)$, где $f(x)$ — непрерывная, положительная функция, выражается формулой

$$F = \int_a^b f(x)\,dx.$$

Приведём два примера применения этой формулы.

Пример 1. Так как уравнение окружности (рис. 28) с центром в начале координат и радиусом R есть

$$x^2+y^2=R^2,$$

то уравнение верхней полуокружности имеет вид

$$y=+\sqrt{R^2-x^2}.$$

Поэтому площадь заштрихованного на чертеже полукруга равна

$$\int_{-R}^{+R}\sqrt{R^2-x^2}\,dx.$$

Полагая $x=R\sin t$, приводим этот интеграл к виду

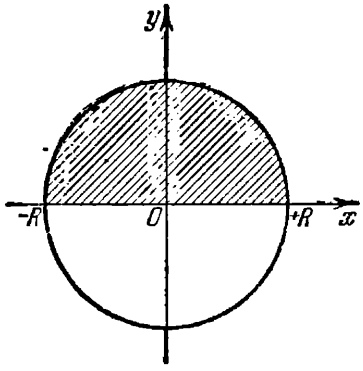


Рис. 28.

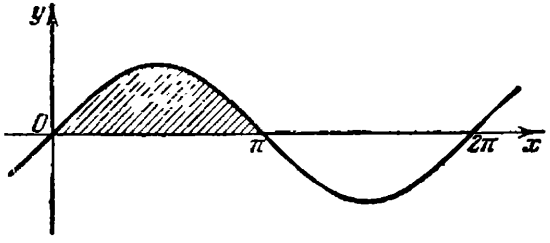


Рис. 29.

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}R^2\cos^2 t\,dt=R^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1+\cos 2t}{2}\,dt=\frac{R^2}{2}\left[t+\frac{1}{2}\sin 2t\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=\frac{\pi R^2}{2}.$$

Поэтому площадь всего круга равна πR^2.

Пример 2. Рассмотрим площадь фигуры, ограниченной осью Ox и полуволной синусоиды $y=\sin x$ (рис. 29). Очевидно, эта площадь равна

$$F=\int_0^{\pi}\sin x\,dx=[-\cos x]_0^{\pi}=2.$$

Любопытно, что она выразилась без каких бы то ни было иррациональностей.

Рассмотрим теперь площадь F фигуры, содержащейся между линиями $x=a$, $x=b$, $y=f(x)$, $y=g(x)$, где $f(x)$ и $g(x)$ — две непрерывные, положительные[1]) функции, заданные на $[a, b]$ и удовлетворяющие неравенству $f(x)<g(x)$ (рис. 30).

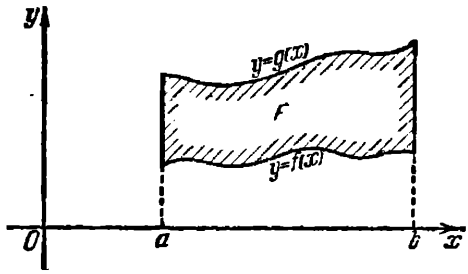


Рис. 30.

Вполне очевидно, что эта площадь выражается формулой

$$F=\int_a^b [g(x)-f(x)]\,dx$$

или, если положить $g(x)-f(x)=r(x)$, формулой

$$F=\int_a^b r(x)\,dx. \qquad (1)$$

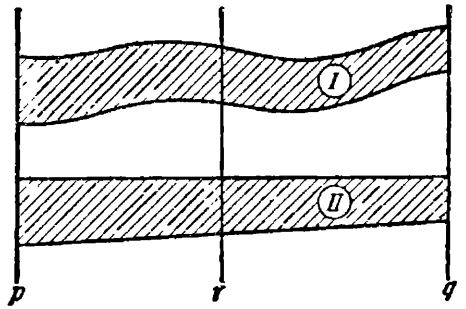


Рис. 31.

Последняя формула показывает, что форма фигуры никакой роли не играет. Важна лишь длина $r(x)$ отрезка ординаты между линиями $y=f(x)$ и $y=g(x)$.

Отсюда следует, что, взяв другую пару функций $f_1(x)$ и $g_1(x)$, подчинённую условию $g_1(x)-f_1(x)=r(x)$, мы получим фигуру, равновеликую прежней. Этот результат, установленный ещё в XVII в. одним из предшественников Ньютона и Лейбница итальянцем Кавальери, допускает и чисто геометрическую формулировку:

Принцип Кавальери. *Если две плоские фигуры I и II содержатся между двумя параллельными прямыми p и q* (рис. 31) *и обладают тем свойством, что в сечении их любой прямой r, параллельной p и q, получаются отрезки одинаковой длины, то эти фигуры имеют одну и ту же площадь.*

Легко показать[2]), что в том случае, когда упомянутые отрезки не равны друг другу, но находятся в некотором постоянном отношении, то в том же отношении будут находиться и площади фигур *I* и *II*.

---

[1]) Легко видеть, что условие положительности $f(x)$ и $g(x)$ можно было бы отбросить.

[2]) В самом деле, одна из площадей будет выражаться формулой (1), а другая — аналогичной формулой $F_1=\int_a^b r_1(x)\,dx$. Если $r_1(x)=k\,r(x)$, то и $F_1=kF$.

29. Вычисление объёмов. Рассмотрим тело T, содержащееся между параллельными плоскостями $x=a$ и $x=b$. Допустим, что в сечении тела T каждой плоскостью $x=x_0$, перпендикулярной к оси Ox, получается фигура $T(x_0)$, имеющая площадь $F(x_0)$, причём $F(x)$ есть непрерывная функция аргумента x (рис. 32).

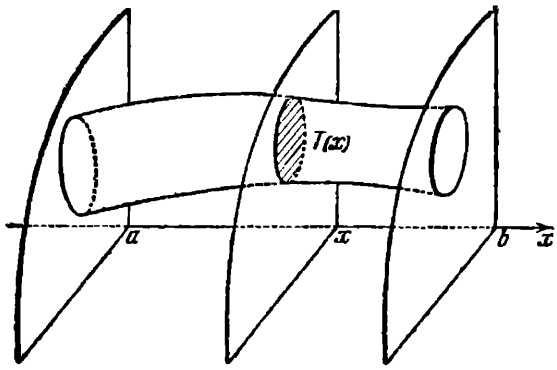


Рис. 32.

Поставим вопрос об объёме тела T. При этом речь здесь должна итти не только о вычислении объёма, но прежде всего о логическом определении этого понятия[1]).

Разобьём промежуток $[a, b]$ точками $x_0=a<x_1<x_2<\ldots<x_n=b$ и проведём плоскости $x=x_k$. Эти плоскости разрежут тело T на n тонких слоёв.

В простых случаях каждый такой слой можно приближённо принять за цилиндр с объёмом

$$F(x_k)(x_{k+1}-x_k)$$

[мы считаем, что прямой цилиндр с основанием, имеющим площадь F, и высотой h имеет объём $V=Fh$]. Поэтому величину суммы

$$\sigma=\sum_{k=0}^{n-1} F(x_k)(x_{k+1}-x_k)$$

естественно считать приближённой мерой объёма V тела T (причём в этот момент рассуждения у нас всё ещё нет точного определения этого понятия!). Но тогда опять-таки естественно принять (что мы и делаем) за самое определение объёма V *предел* суммы σ при $\lambda=\max(x_{k+1}-x_k)\to 0$. Доказывать это определение (как и всякое другое), конечно, не надо. Вместо этого надо подчеркнуть, что упомянутый предел для рассмотренного класса тел всегда существует, ибо σ есть интегральная сумма непрерывной функ-

---

[1]) Развёрнутый анализ понятия объёма будет дан в четвёртой книге «Энциклопедия элементарной математики». Здесь же это понятие рассматривается лишь в связи с использованием интеграла (однократного) для вычисления объёма и общего определения не даётся.

ции $F(x)$. Из этого же обстоятельства немедленно вытекает и формула

$$V=\int_a^b F(x)\,dx, \tag{2}$$

позволяющая вычислять объём тела по площадям его поперечных сечений [1]).

Рассмотрим несколько примеров.

Пример 1. Пусть тело T есть *конус* (рис. 33). Выберем за ось Ox прямую, совпадающую с осью симметрии конуса, направив её от вершины конуса к его основанию и поместив начало O в вершине конуса. Сечение $T(x)$ конуса — круг, радиус которого, очевидно, равен $\frac{R}{H}x$, где R и H соответственно суть радиус основания и высота конуса. Отсюда

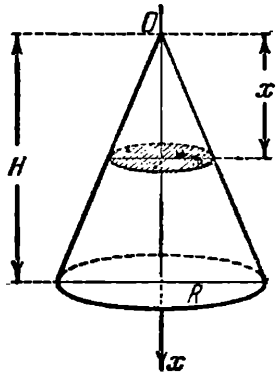


Рис. 33.

$$V=\int_0^H \pi\left(\frac{R}{H}x\right)^2 dx=\frac{1}{3}\pi R^2 H.$$

Пример 2. Столь же просто находится *объём шара*. Действительно, рассмотрим сначала полушар. Введём обозначения, указанные на рис. 34. Радиус r круга $T(x)$, заштрихованного на рис. 34, равен $\sqrt{R^2-x^2}$ (рис. 35). Значит, объём полу-

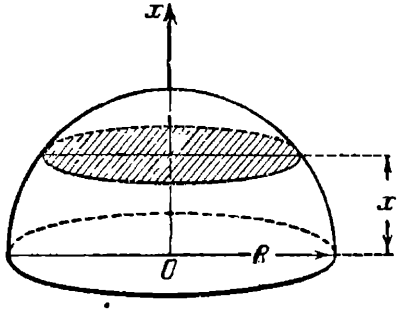


Рис. 34.

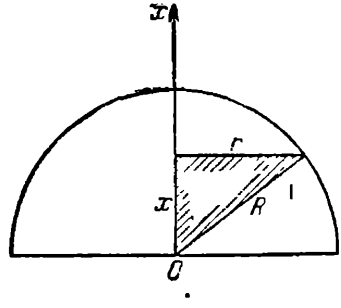


Рис. 35.

[1]) Так как форма сечения $T(x)$ оказалась несущественной, а объём выразился лишь через площадь этого сечения, то из (2) вытекает

Принцип Кавальери для объёмов. *Если два тела I и II содержатся между двумя параллельными плоскостями P и Q и обладают тем свойством, что в сечении их любой плоскостью R, параллельной P и Q, получаются фигуры, имеющие одинаковую площадь, то объёмы тел I и II равны.*

Ясно также, что в том случае, когда упомянутые площади находятся в некотором постоянном отношении (не зависящем от выбора плоскости R), то в том же отношении будут находиться и объёмы тел.

шара выражается формулой [1])

$$\int_0^R \pi (R^2 - x^2)\, dx = \frac{2}{3}\pi R^3,$$

откуда следует и известная формула для объёма шара

$$V = \frac{4}{3}\pi R^3.$$

Следующие примеры ещё ярче демонстрируют чрезвычайную силу общих методов интегрального исчисления.

Пример 3. Пусть T есть тело, отсекаемое от прямого кругового цилиндра плоскостью, проходящей через диаметр его основания (рис. 36). Это тело называется «цилиндрическим отрезком». Положим (мы придерживаемся обозначений рис. 36) $AB = H$, $OA = R$ и выразим объём V тела T через R и H.

Для этого рассмотрим треугольник $O_1A_1B_1$, получающийся в сечении тела T плоскостью, параллельной плоскости OAB и отстоящей от последней на расстояние $OO_1 = x$.

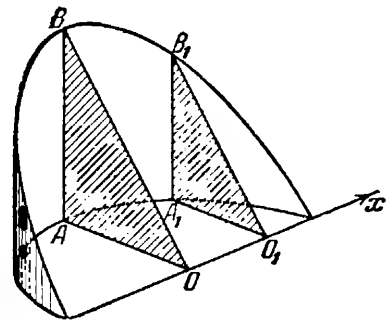


Рис. 36.

Площадь упомянутого треугольника равна $F(x) = \frac{1}{2} O_1A_1 \cdot A_1B_1$.

Из теоремы Пифагора следует, что $O_1A_1 = \sqrt{R^2 - x^2}$, а из подобия треугольников OAB и $O_1A_1B_1$ вытекает, что $A_1B_1 : AB = O_1A_1 : OA$, откуда

$$A_1B_1 = \frac{H}{R}\sqrt{R^2 - x^2}.$$

---

[1]) Эта формула получается также при помощи принципа Кавальери из формул объёмов цилиндра и конуса. Именно, поместим полушар радиуса R между плоскостями верхнего и нижнего оснований цилиндра радиуса R и высоты R, из которого «высверлен» конус того же радиуса и той же высоты (рис. 37). Плоскость H, изображённая на чертеже, пересекает полушар по

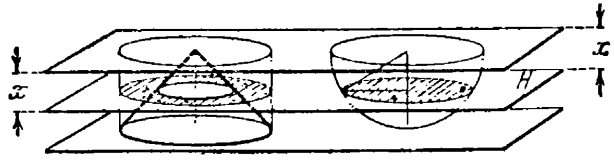


Рис. 37.

кругу радиуса $\sqrt{R^2 - x^2}$, цилиндр — по кругу радиуса R, а конус — по кругу радиуса x. Поэтому площади обоих заштрихованных сечений равны $\pi(R^2 - x^2)$. Значит, объём полушара равен разности $\pi R^3 - \frac{1}{3}\pi R^3 = \frac{2}{3}\pi R^3$. Этот интересный вывод целесообразно сообщать и в средней школе.

Стало быть,

$$F(x)=\frac{H}{2R}(R^2-x^2).$$

Согласно формуле (2) интересующий нас объём есть

$$V=\int_{-R}^{R}\frac{H}{2R}(R^2-x^2)\,dx=\frac{H}{2R}\left[R^2x-\frac{x^3}{3}\right]_{-R}^{+R},$$

откуда

$$V=\frac{2}{3}R^2H.$$

Эту задачу можно решить и иначе. Именно, проведём ось Ox вдоль радиуса OA, т. е. перпендикулярно к диаметру, через который проходит плоскость, отсекающая наш «цилиндрический отрезок» T (рис. 38). Плоскость, перпендикулярная к OA и пересекающая OA в точке P, где $OP=x$, высекает из T прямоугольник $T(x)$, заштрихованный на рис. 38. Его площадь равна

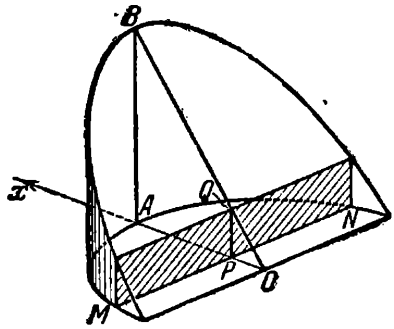


Рис. 38.

$$F(x)=MN\cdot PQ.$$

Но по теореме Пифагора хорда MN равна

$$MN=2\sqrt{R^2-x^2}.$$

Из подобия же треугольников OAB и OPQ следует, что

$$PQ:AB=OP:OA,$$

откуда

$$PQ=\frac{H}{R}x$$

и

$$V=\int_0^R 2\frac{H}{R}x\sqrt{R^2-x^2}\,dx=\frac{H}{R}\int_0^R\sqrt{R^2-x^2}\,d(x^2).$$

При помощи подстановки $R^2-x^2=z$ находим:

$$V=-\frac{H}{R}\int_{R^2}^{0}\sqrt{z}\,dz=-\frac{H}{R}\left[\frac{2}{3}z^{\frac{3}{2}}\right]_{R^2}^{0}$$

и окончательно, как и выше,

$$V=\frac{2}{3}R^2H.$$

Совпадение численных значений V, найденных при двух различных способах решения последнего примера, приводит к некоторым общим соображениям принципиального характера. Именно, возвращаясь к основной формуле (2)

$$V = \int_a^b F(x)\,dx, \tag{2}$$

мы видим, что для вычисления объёма по этой формуле надо начать с выбора оси Ox. При другом выборе этой оси изменятся сечения $T(x)$ и их площади $F(x)$. Но тогда естественно возникает сомнение, не изменится ли при этом и величина объёма, вычисляемого по формуле (2). Разумеется, если бы это было так, т. е. если бы объём тела зависел от столь субъективного момента, как выбор оси Ox, то это сильно понизило бы ценность данного нами определения объёма. Поэтому указанное определение следовало бы сопровождать доказательством независимости величины объёма от выбора оси Ox. Мы, однако, позволим себе оставить этот вопрос в стороне, ибо упомянутое доказательство не только было бы весьма сложным, но и потребовало бы существенного сужения класса рассматриваемых тел. Кроме того, весь вопрос относится не столько к интегральному исчислению, сколько к области измерительной геометрии.

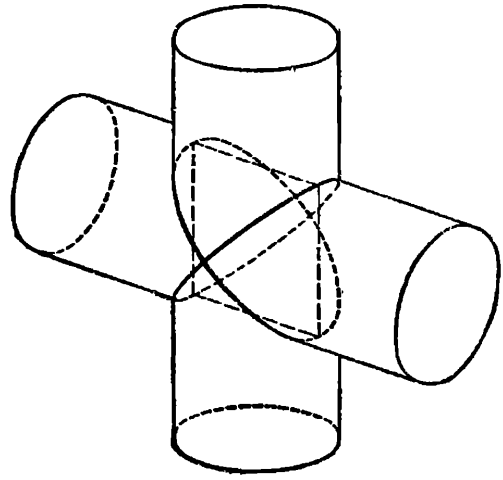


Рис. 39.

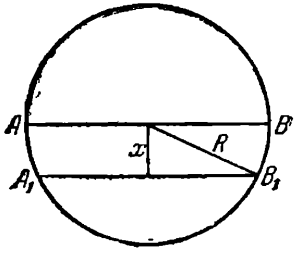


Рис. 40.

Пример 4. Найдём объём V тела T, являющегося общей частью двух одинаковых цилиндров, оси которых пересекаются под прямым углом (рис. 39). Несмотря на некоторую затруднительность отчётливого представления этого тела, поставленная задача решается просто. Именно, назовём «осевой плоскостью» плоскость, содержащую оси обоих цилиндров. Плоскость, параллельная осевой и отстоящая от неё на расстояние x, пересечёт каждый из наших цилиндров по прямоугольной полосе, ширина которой, как видно из рис. 40, равна

$$A_1B_1 = 2\sqrt{R^2 - x^2}.$$

В таком случае сечение $T(x)$ тела T указанной плоскостью оказывается квадратом со стороной, равной A_1B_1. Отсюда ясно, что

$$V=\int_{-R}^{+R} 4(R^2-x^2)\,dx=4\left[R^2x-\frac{x^3}{3}\right]_{-R}^{+R}$$

т. е.

$$V=\frac{16}{3}R^3.$$

Остановимся на одном важном частном случае общей формулы (2). Представим себе непрерывную кривую $y=f(x)$, где $a\leqslant x\leqslant b$. Допустим, что эта кривая вращается вокруг оси Ox, и рассмотрим тело T, ограниченное поверхностью вращения упомянутой кривой и плоскостями $x=a$ и $x=b$ (рис. 41). В нашем случае сечения $T(x)$ суть круги радиуса $f(x)$, и потому объём тела вращения T будет

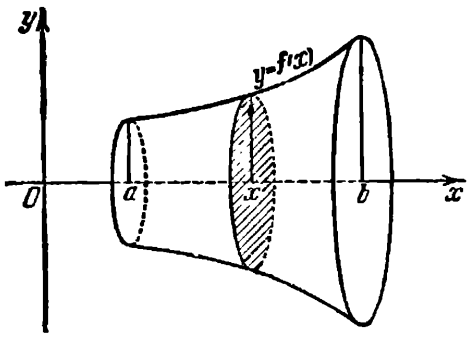


Рис. 41.

$$V=\pi\int_a^b f^2(x)\,dx. \qquad (3)$$

Найдём, например, объём тела, ограниченного поверхностью вра-

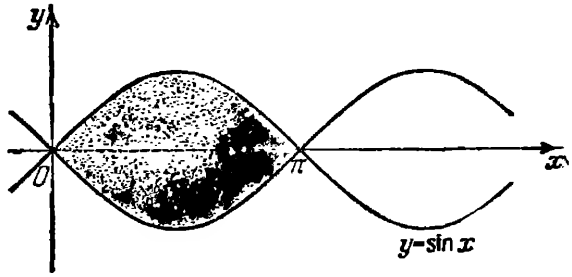


Рис. 42.

щения полуволны синусоиды (рис. 42). Согласно формуле (3)

$$V=\pi\int_0^{\pi}\sin^2 x\,dx=\pi\int_0^{\pi}\frac{1-\cos 2x}{2}\,dx,$$

и стало быть,

$$V=\frac{\pi}{2}\left[x-\frac{1}{2}\sin 2x\right]_0^{\pi}=\frac{\pi^2}{2}.$$

30. Длина дуги кривой. Рассмотрим кривую $y=f(x)$, где $f(x)$ — непрерывная функция, заданная на промежутке $[a, b]$. Разделим $[a, b]$ точками $x_0=a<x_1<\ldots<x_n=b$ и образуем ломаную с вершинами $M_k(x_k, f(x_k))$ $(k=0, 1, \ldots, n)$. Если существует конечный предел s длины этой ломаной, когда наибольшее из её звеньев стремится к нулю, то он называется *длиной* нашей кривой.

Вообще говоря, не у всякой непрерывной кривой существует длина (можно доказать, что упомянутый предел существует всегда, но он может оказаться бесконечным). Те кривые, у которых длина существует, называются *спрямляемыми*.

Теорема. *Если у функции $f(x)$ существует непрерывная производная $f'(x)$, то кривая $y=f(x)$ спрямляема, и её длина равна*

$$s=\int_a^b \sqrt{1+f'^2(x)}\,dx. \tag{4}$$

Действительно, длина одного звена M_kM_{k+1} упомянутой выше ломаной равна, очевидно,

$$M_kM_{k+1}=\sqrt{(x_{k+1}-x_k)^2+[f(x_{k+1})-f(x_k)]^2}.$$

Но по формуле Лагранжа

$$f(x_{k+1})-f(x_k)=f'(\xi_k)(x_{k+1}-x_k),$$

где $x_k<\xi_k<x_{k+1}$. Таким образом, длина σ всей ломаной такова:

$$\sigma=\sum_{k=0}^{n-1}\sqrt{1+f'^2(\xi_k)}\,(x_{k+1}-x_k).$$

Так как это выражение есть интегральная сумма непрерывной функции

$$\sqrt{1+f'^2(x)},$$

то [при стремлении к нулю величины $\lambda=\max(x_{k+1}-x_k)$] оно стремится к интегралу, стоящему в (4). Если обозначить через μ наибольшее из звеньев M_kM_{k+1}, то, очевидно, будет $\lambda\leqslant\mu$. Поэтому когда стремится к нулю μ, то λ и подавно стремится к нулю, откуда и вытекает справедливость теоремы.

Пример. Рассмотрим окружность радиуса R. Помещая её центр в начало координат, запишем её уравнение в виде

$$x^2+y^2=R^2.$$

Тогда уравнение верхней полуокружности будет

$$y=\sqrt{R^2-x^2},$$

откуда следует, что

$$y'=\frac{-x}{\sqrt{R^2-x^2}}.$$

Если рассматривать всю верхнюю полуокружность, т. е. изменять x в промежутке $[-R, +R]$, то мы не сможем применить доказанную выше теорему, ибо производная y' не существует при $x=\pm R$. Поэтому мы рассмотрим лишь ту половину верхней полуокружности, которая лежит между прямыми $y=x$ и $y=-x$ (т. е. дугу AB, рис. 43).

Так как абсциссы точек A и B соответственно равны $-\frac{R\sqrt{2}}{2}$ и $+\frac{R\sqrt{2}}{2}$, а в промежутке $\left[-\frac{R\sqrt{2}}{2}, +\frac{R\sqrt{2}}{2}\right]$ функция $y=\sqrt{R^2-x^2}$ уже удовлетворяет условиям теоремы, то дуга AB спрямляема и её длина такова:

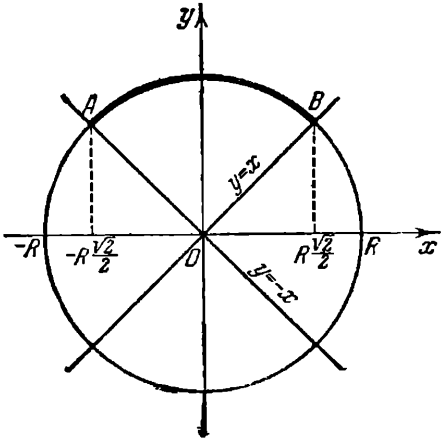


Рис. 43.

$$s_{AB}=\int\limits_{-\frac{R\sqrt{2}}{2}}^{+\frac{R\sqrt{2}}{2}}\sqrt{1+y'^2}\,dx=$$

$$=\int\limits_{-\frac{R\sqrt{2}}{2}}^{+\frac{R\sqrt{2}}{2}}\frac{R\,dx}{\sqrt{R^2-x^2}}.$$

Отсюда

$$s_{AB}=R\left[\arcsin\frac{x}{R}\right]_{-\frac{R\sqrt{2}}{2}}^{+\frac{R\sqrt{2}}{2}}=\frac{\pi}{2}R,$$

а так как AB есть в точности одна четверть всей окружности, то окружность спрямляема и её длина равна $2\pi R$.

31. Площадь поверхности вращения. Рассмотрим, как и выше, кривую $y=f(x)$, где $f(x)$ — непрерывная функция, заданная в промежутке $[a, b]$. Для простоты эту функцию мы будем предполагать положительной. Как и выше, впишем в нашу кривую ломаную с вершинами $M_k(x_k, f(x_k))$, где $x_0=a<x_1<x_2<\ldots<x_n=b$. Пусть наша кривая, а вместе с ней и упомянутая ломаная вращаются вокруг оси Ox. Тогда кривая опишет некоторую поверхность вращения (L), а ломаная — вписанную в неё поверхность (L'), составленную из n усечённых конусов (в частном случае вырождающихся в цилиндры). Если при стремлении к нулю наибольшей из хорд M_kM_{k+1} площадь[1]) L' поверхности (L') стремится к конеч-

[1]) Мы считаем известным, что боковая поверхность усечённого конуса имеет площадь, равную его образующей, умноженной на периметр среднего сечения.

ному пределу L, то последний называется *площадью поверхности* (L) [1]).

Теорема. *Если у функции* $f(x)$ *существует непрерывная производная* $f'(x)$, *то поверхность* (L) *имеет площадь, выражающуюся формулой*

$$L = 2\pi \int_a^b f(x)\sqrt{1+f'^2(x)}\,dx. \tag{5}$$

В самом деле, площадь боковой поверхности конуса, образованного вращением звена M_kM_{k+1}, равна

$$\pi[f(x_k)+f(x_{k+1})]\sqrt{(x_{k+1}-x_k)^2+[f(x_{k+1})-f(x_k)]^2}.$$

Но по формуле Лагранжа

$$f(x_{k+1})-f(x_k)=f'(\xi_k)(x_{k+1}-x_k) \quad (x_k<\xi_k<x_{k+1}).$$

Таким образом, сохраняя вышеуказанные обозначения, будем иметь

$$L'=\pi\sum_{k=0}^{n-1}[f(x_k)+f(x_{k+1})]\sqrt{1+f'^2(\xi_k)}\,(x_{k+1}-x_k).$$

Это выражение сходно с суммой

$$\sigma=2\pi\sum_{k=0}^{n-1}f(\xi_k)\sqrt{1+f'^2(\xi_k)}\,(x_{k+1}-x_k),$$

которая стремится при $\lambda=\max(x_{k+1}-x_k)\to 0$ к интегралу, стоящему в формуле (5) [ибо σ есть интегральная сумма непрерывной функции $f(x)\sqrt{1+f'^2(x)}$]. Так как из стремления к нулю наибольшей хорды M_kM_{k+1} вытекает стремление к нулю и величины λ, то достаточно удостовериться в том, что

$$\lim_{\lambda\to 0}(L'-\sigma)=0. \tag{6}$$

С этой целью, взяв произвольное $\varepsilon>0$, найдём столь малое $\delta>0$, чтобы из неравенства $|x'-x''|<\delta$ следовало неравенство $|f(x')-f(x'')|<\varepsilon$. Пусть $\lambda<\delta$. Тогда каждая из разностей $|f(x_k)-f(\xi_k)|$ и $|f(x_{k+1})-f(\xi_k)|$ будет меньше, чем ε. Но

$$L'-\sigma=$$
$$=\pi\sum_{k=0}^{n-1}\{[f(x_k)-f(\xi_k)]+[f(x_{k+1})-f(\xi_k)]\}\sqrt{1+f'^2(\xi_k)}\,(x_{k+1}-x_k).$$

---

[1]) Это определение площади поверхности пригодно только для поверхностей вращения. В общем случае определение площади поверхности значительно усложняется (см. об этом в четвёртой книге Э. э. м.).

Следовательно, обозначая через M наибольшее значение непрерывной функции $\sqrt{1+f'^2(x)}$, мы будем для $\lambda<\delta$ иметь, что

$$|L'-\sigma|\leqslant 2M\pi\varepsilon\sum_{k=0}^{n-1}(x_{k+1}-x_k)=2M\pi(b-a)\varepsilon.$$

Так как правая часть этого неравенства сколь угодно мала вместе с ε, то доказано соотношение (6), а с ним и теорема.

В качестве примера рассмотрим вопрос о площади *шарового пояса.*

Пусть полуокружность

$$y=\sqrt{R^2-x^2}$$

вращается вокруг оси Ox. Рассмотрим площадь поверхности, образованной вращением той дуги этой полуокружности, которая лежит между прямыми $x=-h$ и $x=+h$, где [1]) $0<h<R$. У нас

$$y'=\frac{-x}{\sqrt{R^2-x^2}},\quad \sqrt{1+y'^2}=\frac{R}{\sqrt{R^2-x^2}},$$

откуда

$$L=2\pi\int_{-h}^{+h}\sqrt{R^2-x^2}\cdot\frac{R}{\sqrt{R^2-x^2}}\,dx=2\pi R\int_{-h}^{+h}dx=4\pi Rh.$$

Такова площадь шарового пояса. Устремляя h к R и переходя к пределу, мы получим [2]) площадь *поверхности шара* $L=4\pi R^2$.

32. Общие указания по поводу приложений интегрального исчисления и его связей с дифференциальным исчислением. Большая часть приложений теории определённого интеграла строится по одной и той же схеме. Здесь мы постараемся выявить характерные черты этой важной схемы.

Пусть всякому промежутку $[\alpha, \beta]$, содержащемуся в некотором закреплённом промежутке $[a, b]$, отвечает значение определённой физической или геометрической величины P, которое мы будем обозначать через

$$P([\alpha, \beta]).$$

Такую величину естественно называть «функцией промежутка $[\alpha, \beta]$».

---

[1]) Применить теорему п° 31 для всего промежутка $[-R, +R]$ здесь нельзя, так как функция $y=\sqrt{R^2-x^2}$ не имеет производной при $x=\pm R$.

[2]) Последнее заключение не вполне обосновано. Рассматривать площадь поверхности шара как предел площади шарового пояса, строго говоря, мы не вправе, ибо по определению должны понимать её только как предел площади L' поверхности (L'). Можно было бы, непосредственно изучая упомянутую площадь L' для данного примера, доказать формулу $L=4\pi R^2$, однако это было бы довольно громоздко, и мы задерживаться на этом не будем.

Рассмотрим, например, непрерывную положительную функцию $f(x)$, заданную для $x \in [a, b]$. Тогда с каждым промежутком $[\alpha, \beta]$, содержащимся в $[a, b]$, мы связываем величину $F([\alpha, \beta])$ площади криволинейной трапеции, ограниченной линиями $y=0$, $y=f(x)$, $x=\alpha$, $x=\beta$. Площадь $F([\alpha, \beta])$ и будет функцией промежутка $[\alpha, \beta]$. Другим примером может служить объём $V([\alpha, \beta])$ тела, образованного вращением указанной трапеции вокруг оси Ox. Вот ещё один простой пример: пусть на $[a, b]$ непрерывным образом распределена масса; тогда количество $m([\alpha, \beta])$ массы, попавшее на промежуток $[\alpha, \beta]$, будет функцией этого промежутка.

Функция промежутка $P([\alpha, \beta])$ называется *аддитивной*, если при $\alpha < \gamma < \beta$ будет

$$P([\alpha, \beta]) = P([\alpha, \gamma]) + P([\gamma, \beta]).$$

Функции $F([\alpha, \beta])$, $V([\alpha, \beta])$, $m([\alpha, \beta])$, о которых мы говорили только что, очевидно, аддитивны.

Рассмотрим аддитивную функцию промежутка $P([\alpha, \beta])$ и допустим, что на основном промежутке $[a, b]$ определена непрерывная функция $p(x)$, связанная с $P([\alpha, \beta])$ следующим соотношением

$$P([x, x+\Delta x]) = p(x)\,\Delta x + \rho([x, x+\Delta x]), \tag{7}$$

где $\rho([x, x+\Delta x])$ есть функция, обладающая тем свойством, что[1])

$$\lim_{\Delta x \to 0} \frac{\rho([x, x+\Delta x])}{\Delta x} = 0. \tag{8}$$

Грубо говоря, соотношение (7) означает, что значение величины P, отвечающее весьма малому промежутку $[x, x+\Delta x]$, почти пропорционально его длине Δx, ибо величина $\rho([x, x+\Delta x])$ (при бесконечно малом Δx) есть бесконечно малая высшего порядка.

Покажем, что в этих условиях значение $P([a, b])$ величины P, отвечающее всему промежутку $[a, b]$, выражается формулой

$$P([a, b]) = \int_a^b p(x)\,dx. \tag{9}$$

В самом деле, разбив $[a, b]$ точками $x_0 = a < x_1 < \ldots < x_n = b$, на основании аддитивности величины P получим

$$P([a, b]) = \sum_{k=0}^{n-1} P([x_k, x_{k+1}]).$$

---

[1]) Точный смысл равенства (8) таков: всякому $\varepsilon > 0$ отвечает такое $\delta > 0$ (не зависящее от x), что при $0 < \Delta x < \delta$ будет $\left| \frac{\rho([x, x+\Delta x])}{\Delta x} \right| < \varepsilon$,

Отсюда и из (7) вытекает, что

$$P([a,\ b]) = \sum_{k=0}^{n-1} p(x_k)(x_{k+1} - x_k) + \sum_{k=0}^{n-1} \rho([x_k,\ x_{k+1}]). \qquad (10)$$

При измельчении дробления первая написанная здесь сумма стремится к интегралу, стоящему в (9). Остаётся показать, что вторая сумма стремится к нулю. С этой целью возьмём $\varepsilon > 0$. Согласно (8), этому ε отвечает такое $\delta > 0$, что при $0 < \Delta x < \delta$ будет

$$|\rho([x,\ x + \Delta x])| < \varepsilon \Delta x.$$

Найдя подобное δ, рассмотрим такое дробление $[a,\ b]$, у которого $\lambda = \max(x_{k+1} - x_k) < \delta$. Тогда

$$\left|\sum_{k=0}^{n-1} \rho([x_k,\ x_{k+1}])\right| < \varepsilon \sum_{k=0}^{n-1} (x_{k+1} - x_k) = \varepsilon(b - a).$$

Таким образом, вторая сумма правой части (10) действительно стремится к нулю, чем и доказано (9).

Итак, если нам удастся установить «приближённую пропорциональность» величины $P([x,\ x + \Delta x])$ и длины Δx малого промежутка $[x,\ x + \Delta x]$, то уже отсюда будет вытекать возможность вычислять величину P по формуле (9).

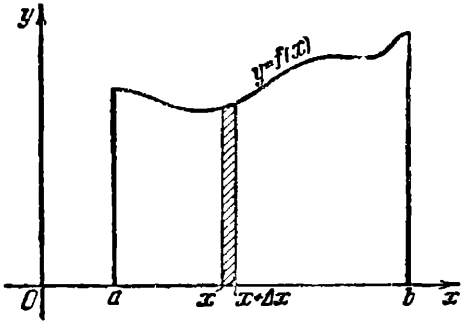


Рис. 44.

В этом и состоит вышеупомянутая общая схема для приложений интегрального исчисления. Именно, всё сводится к приближённому (с точностью до малых высшего порядка) выделению из «элементарного» слагаемого $P([x,\ x + \Delta x])$ величины вида $p(x)\Delta x$.

Так, например, полоску, изображённую на рис. 44, приближённо можно принять за прямоугольник с основанием Δx и высотой $f(x)$. Отсюда сразу вытекает, что

$$F([a,\ b]) = \int_a^b f(x)\,dx.$$

Покажем применение указанной схемы ещё на одном примере. Вообразим себе канал, заполненный водой и запертый щитом, изображённым на рис. 45.

Пусть свободная поверхность воды находится на уровне AB, а глубина канала равна h. Поставим вопрос о величине P давления, испытываемого щитом со стороны воды.

Обозначим через $l(x)$ длину горизонтальной прямой, проведённой на щите на расстоянии x от AB. Давление $P([x, x+\Delta x])$, испытываемое полоской, содержащейся между горизонтальными прямыми, отстоящими от AB на расстояния x и $x+\Delta x$ (эта полоска заштрихована на рис. 45), мы можем подсчитать приближённо, если, во-первых, примем полоску за прямоугольник с основанием $l(x)$ и высотой Δx, а во-вторых, будем считать, что все точки полоски находятся на одной и той же глубине x. Именно, применяя известное правило гидростатики, гласящее, что давление воды на погружённую в неё малую площадку равно площади этой площадки, умноженной на глубину её погружения, мы получаем, что [с точностью до бесконечно малой порядка высшего чем Δx] будет

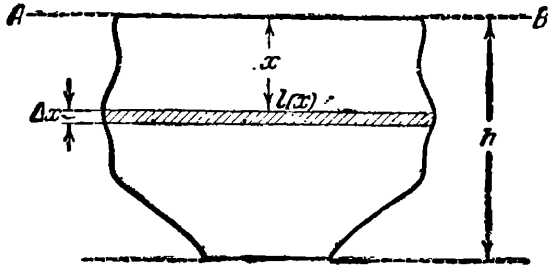


Рис. 45.

$$P([x, x+\Delta x]) = x\,l(x)\,\Delta x.$$

На основании сказанного выше можем сразу написать формулу для всего давления

$$P = \int_0^h x\,l(x)\,dx.$$

Задача решена в общем виде. Для получения численного значения P надо лишь задать функцию $l(x)$, которая определяется формой щита.

Отметим одну интересную в методологическом отношении деталь. Очень часто предыдущие соображения показывают нам не только, как вычислить величину P, но и открывают путь для её формального определения. По существу именно таким путём мы и шли, определяя понятие площади в п° 28 и понятие объёма в п° 29. Таким образом, интегральное исчисление не только даёт возможность вычислять различные величины, но одновременно вооружает нас и неким общим методом конструирования определений этих величин.

Соотношение (7) позволяет взглянуть на связь между функцией промежутка $P([\alpha, \beta])$ и функцией точки $p(x)$ ещё и с другой стороны. Именно, из этого соотношения сразу вытекает, что

$$p(x) = \lim_{\Delta x \to 0} \frac{P([x, x+\Delta x])}{\Delta x}. \qquad (11)$$

Нетрудно понять, что предельный переход, указываемый в формуле (11), является по существу операцией дифференцирования[1]).

Таким образом, если переход от функции точки $p(x)$ к функции промежутка $P([\alpha, \beta])$ производится при помощи интегрирования по формуле (9), то обратный переход осуществляется при помощи дифференцирования по формуле (11). Здесь взаимно обратный характер операций дифференцирования и интегрирования выступает весьма наглядно и притом в очень общей форме.

Характеристика какого-либо явления при помощи функции точки представляет собой локальную характеристику этого явления, то-есть характеристику в малом, а характеристика при помощи функции промежутка — характеристику в целом, интегральную характеристику. Поэтому мы можем формулировать задачу интегрирования, как задачу установления свойств явления в целом на основании его локальных свойств, а задачу дифференцирования, — как задачу локальной характеристики явления на основании его свойств в целом. Такое взаимоотношение между задачами интегрирования и дифференцирования сохраняется и при самых разнообразных их обобщениях.

---

[1]) Действительно, полагая $F(x) = P([a, x])$, мы сможем соотношение

$$P([a, x+\Delta x]) = P([a, x]) + P([x, x+\Delta x]),$$

справедливое в силу аддитивности величины P, записать в форме

$$P([x, x+\Delta x]) = F(x+\Delta x) - F(x).$$

Отсюда и из (11) следует, что $p(x) = F'(x)$.

ГЛАВА III

РЯДЫ

§ 7. Ряды с постоянными членами

33. Основные понятия. Характерной чертой математики является тесное слияние двух сторон этой науки — абстрактно-логической и вычислительной. Благодаря своим мощным вычислительным методам, математика является важнейшим орудием современного естествознания и техники. В то же время создание этих методов возможно лишь на основе глубоко разработанных теоретических построений. Теория рядов представляет собой ту область математического анализа, в которой отмеченное слияние наблюдается особенно ярко. Будучи чрезвычайно насыщенной глубоким и тонким логическим материалом, теория рядов является вместе с тем настоящей лабораторией вычислительных методов математики. Достаточно указать хотя бы на то обстоятельство, что именно теория рядов даёт нам средства для составления логарифмических таблиц, таблиц тригонометрических функций, позволяет находить важнейшие постоянные, как, например, число π, и т. п.

Сказанным объясняется то центральное место, которое занимает теория рядов в современном математическом анализе, хотя сами по себе бесконечные ряды, являющиеся предметом изучения этой теории, и представляют, казалось бы, выражения весьма частного вида.

Определение. *Рядом называется выражение вида*

$$a_1 + a_2 + a_3 + \dots, \tag{1}$$

в котором $a_1, a_2, a_3, \dots, a_n, \dots$ (члены ряда) *суть определённые числа, закон построения которых известен.*

Иногда ряд (1) записывают в форме

$$\sum_{k=1}^{\infty} a_k. \tag{2}$$

Самой важной стороной дела при образовании выражения вида (1) является то многоточие, которое поставлено в конце этого выражения. Оно показывает, что множество чисел a_k, участвующих

в определении ряда (1), обязательно бесконечно. Таким образом, с чисто формальной точки зрения, ряды — это суммы, содержащие бесконечное множество слагаемых.

Первый вопрос, возникающий при рассмотрении подобных выражений, заключается в том, имеют ли они какое-либо числовое значение. Оказывается, что приписать разумным образом такое значение удаётся далеко не всем, а лишь так называемым сходящимся рядам [1]).

Пусть дан ряд (1). Образуем последовательность чисел S_1, S_2, S_3, ..., полагая

$$S_n = a_1 + a_2 + \ldots + a_n \qquad (n = 1, 2, 3\ldots).$$

Эти числа называются *частичными суммами* ряда (1).

Если существует конечный предел

$$\lim_{n\to\infty} S_n, \tag{3}$$

то говорят, что ряд (1) *сходится*, а предел (3) называется его *суммой*. Если же предела (3) не существует или он бесконечен, то ряд (1) называется *расходящимся*.

Таким образом, *сумма ряда — это (конечный) предел последовательности его частичных сумм.*

Если ряд (1) имеет сумму S, то пишут

$$S = a_1 + a_2 + a_3 + \ldots,$$

или

$$S = \sum_{k=1}^{\infty} a_k.$$

Если же ряд расходится, то [1]) ему не приписывают никакого числового значения.

Приведём несколько примеров, разъясняющих введённые определения.

Пример 1. Рассмотрим ряд

$$0 + 0 + 0 + 0 + \ldots \tag{4}$$

Здесь $S_n = 0$, а потому и $\lim\limits_{n\to\infty} S_n = 0$, т. е. ряд (4) сходится и сумма его равна нулю.

Пример 2. Рассмотрим ряд

$$1 + 1 + 1 + 1 + \ldots \tag{5}$$

Здесь $S_n = n$, поэтому $\lim\limits_{n\to\infty} S_n = +\infty$ и ряд расходится.

[1]) В более высоких частях теории оказывается возможным и некоторым расходящимся рядам придавать числовое значение, но мы здесь этого делать не будем.

Пример 3. Рассмотрим ряд

$$1+(-1)+1+(-1)+\ldots \tag{6}$$

Здесь $S_n=1$, если n — число нечётное, и $S_n=0$, если n — число чётное. Значит, ряд (6) расходится.

Этот пример весьма поучителен. Он показывает, что нельзя приписывать чрезмерно большое значение аналитическим аппаратам, считая, что если написано некоторое аналитическое выражение, то оно обязательно должно иметь какой-то смысл, и вопрос состоит лишь в том, чтобы доискаться этого смысла. Нет надобности говорить, что последняя концепция неправильна [1]), и педагогически полезны примеры, опровергающие её. Ряд (6) хорошо приспособлен для этой цели. Один математик XVIII в. по поводу этого ряда «рассуждал» так: пусть

$$1+(-1)+1+(-1)+\ldots=S. \tag{7}$$

Тогда, объединяя члены попарно, получим

$$S=[1+(-1)]+[1+(-1)]+[1+(-1)]+\ldots=0+0+0+\ldots=0.$$

С другой стороны, S можно записать и так:

$$S=1+[(-1)+1]+[(-1)+1]+\ldots=1+0+0\ldots=1.$$

Отсюда «следует», что $0=1$. Будучи мистиком, этот математик выводил далее из своего «открытия», что «ничто» (т. е. 0) — это то же самое, что «нечто» (т. е. 1), а потому ничего абсурдного в том, что бог создал мир из ничего, якобы нет.

Приведённые «рассуждения», конечно, лишены какого бы то ни было научного значения, но они характерны как образец той «фетишизации аппарата», о которой говорилось выше. Читателю ясно, что причиной нелепого заключения, что $0=1$, является пренебрежение вопросом о сходимости рассматриваемого ряда. Начиная с первой четверти XIX в., учение о рядах было поставлено на твёрдую почву, благодаря введению точных определений, и все подобные недоразумения отпали.

Пример 4. Рассмотрим ряд

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\ldots \tag{8}$$

Замечая, что

$$\frac{1}{k(k+1)}=\frac{1}{k}-\frac{1}{k+1},$$

мы можем представить частичную сумму S_n так:

$$S_n=\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\ldots+\left(\frac{1}{n}-\frac{1}{n+1}\right).$$

[1]) Хотя иногда она даёт толчок к весьма ценным открытиям (например, мнимые числа, операторное исчисление). Иными словами, эта концепция, будучи логически несостоятельной, представляет полезное эвристическое средство.

Отсюда ясно, что

$$S_n = 1 - \frac{1}{n+1},$$

а так как это выражение при $n \to \infty$ стремится к 1, то ряд (8) сходится и сумма его есть 1.

Пример 5. Большое теоретическое значение имеет так называемый *гармонический* ряд, т. е. ряд

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots \qquad (9)$$

Покажем, что этот ряд расходится. В самом деле, его частичную сумму S_{2^m} можно записать так:

$$S_{2^m} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \\ + \ldots + \left(\frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \ldots + \frac{1}{2^m}\right).$$

Сумма, стоящая в каждых скобках, больше, чем $\frac{1}{2}$, ибо

$$\frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = \frac{1}{4} \cdot 2 = \frac{1}{2},$$

$$\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{8} \cdot 4 = \frac{1}{2},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \ldots + \frac{1}{2^m} > \frac{1}{2^m} \cdot 2^{m-1} = \frac{1}{2}.$$

Отсюда следует, что

$$S_{2^m} > 1 + \frac{m}{2}.$$

Значит, суммы S_n не могут стремиться ни к какому конечному пределу, и ряд (9) расходится.

Пример 6. В качестве последнего примера рассмотрим *геометрическую прогрессию*

$$a + aq + aq^2 + aq^3 + \ldots \qquad (a \neq 0). \qquad (10)$$

Если $q = 1$, то частичная сумма прогрессии имеет вид $S_n = an$ и прогрессия расходится аналогично ряду (5). Если же $q \neq 1$, то по известной формуле алгебры

$$S_n = \frac{a - aq^n}{1 - q}$$

Если $|q| < 1$, то при возрастающем n величина q^n стремится к нулю и

$$S_n \to \frac{a}{1-q}.$$

Если же $|q|>1$, то q^n стремится к бесконечности, и у S_n конечного предела нет.

Если, наконец, $|q|=1$, то, останавливаясь на случае $q=-1$ (ибо случай $q=+1$ уже рассмотрен), приходим к прогрессии вида

$$a+(-a)+a+(-a)+\ldots,$$

которая аналогично ряду (6) оказывается расходящейся.

Таким образом *геометрическая прогрессия* (10) *сходится тогда и только тогда, когда абсолютная величина её знаменателя меньше единицы*

$$|q|<1.$$

В последнем случае сумма прогрессии равна

$$S=\frac{a}{1-q}.$$

34. Простейшие свойства рядов.

Теорема 1. *Если члены сходящегося ряда, не меняя их порядка, объединить в конечные группы и составить ряд из сумм этих групп, то он будет сходиться и иметь сумму, равную сумме первоначального ряда.*

Иначе говоря, если

$$a_1+a_2+a_3+\ldots=S \tag{11}$$

и $n_1<n_2<n_3<\ldots$, то

$$(a_1+a_2+\ldots+a_{n_1})+(a_{n_1+1}+\ldots+a_{n_2})+ \\ +(a_{n_2+1}+\ldots+a_{n_3})+\ldots=S. \tag{12}$$

В самом деле, если S_n и S_n^* суть соответственно частичные суммы рядов (11) и (12), то, как легко видеть,

$$S_k^*=S_{n_k}.$$

Это означает, что последовательность S_1^*, S_2^*, S_3^*, ... является частичной по отношению к последовательности S_1, S_2, S_3, ... Так как последняя сходится к S, то это верно и относительно последовательности $\{S_k^*\}$.

Доказанную теорему можно коротко формулировать так: *члены сходящегося ряда можно заключать в скобки*. Ещё иначе можно сказать, что сходящиеся ряды обладают «сочетательным свойством».

Интересно отметить, что из сходимости ряда (12) не вытекает сходимость ряда (11). Это видно хотя бы из следующего примера: хотя ряд

$$[1+(-1)]+[1+(-1)]+[1+(-1)]+\ldots$$

сходится, но ряд

$$1+(-1)+1+(-1)+1+\ldots$$

расходится. Таким образом, «раскрывать скобки» можно не всегда.

Теорема 2. *Если все члены сходящегося ряда умножить на одно и то же число, то вновь полученный ряд будет сходиться, и его сумма будет равна сумме первоначального ряда, умноженной на то же число.*

Иначе говоря, если

$$a_1 + a_2 + a_3 + \ldots = S, \tag{13}$$

то

$$ca_1 + ca_2 + ca_3 + \ldots = cS. \tag{14}$$

Для доказательства вводим частичные суммы S_n и S_n^* рядов (13) и (14). Из соотношений

$$S_n^* = cS_n, \quad S_n \to S$$

вытекает, что $S_n^* \to cS$, чем и доказана теорема.

Теорема 3. *Сходящиеся ряды можно почленно складывать.*

Более подробно эта теорема может быть сформулирована так: если

$$a_1 + a_2 + a_3 + \ldots = A, \tag{15}$$

$$b_1 + b_2 + b_3 + \ldots = B, \tag{16}$$

то

$$(a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) + \ldots = A + B. \tag{17}$$

Для доказательства теоремы обозначим соответственно через A_n, B_n и S_n частичные суммы рядов (15), (16) и (17). Тогда

$$S_n = A_n + B_n, \quad A_n \to A, \quad B_n \to B,$$

откуда следует, что $S_n \to A + B$.

Введём важное понятие остатка ряда. Именно, если

$$a_1 + a_2 + a_3 + a_4 + \ldots \tag{18}$$

есть некоторый ряд, а m — какое-нибудь натуральное число, то *остатком ряда* (8) *после m-го члена* называется ряд

$$a_{m+1} + a_{m+2} + a_{m+3} + \ldots \tag{19}$$

Теорема 4. *Сам ряд и его остаток сходятся или расходятся одновременно.*

В самом деле, пусть частичные суммы рядов (18) и (19) суть соответственно S_n и S_n^*. Тогда при $n > m$

$$S_n = S_m + S_{n-m}^*. \tag{20}$$

Допустим, что остаток (19) сходится и его сумма равна R_m. Тогда при стремлении n к бесконечности будет

$$S_{n-m}^* \to R_m,$$

откуда следует, что $S_n \to S_m + R_m$. Иначе говоря, ряд (18) также сходится, и его сумма есть

$$S = S_m + R_m. \tag{21}$$

Обратно, если сходится ряд (18) и сумма его равна S, то из того же соотношения (20) вытекает, что

$$\lim_{n\to\infty} S_{n-m}^{*} = S - S_m,$$

чем и доказана сходимость ряда (19).

Если ряд сходится, то, как это видно из равенства (21), сумма его остатка после m-го члена в точности равна разности между суммой S всего ряда и его частичной суммой S_m. Так как при возрастании m частичная сумма S_m будет стремиться к S, то отсюда следует

Теорема 5. *Сумма остатка сходящегося ряда после m-го члена стремится к нулю при возрастании m*

$$\lim_{m\to\infty} (a_{m+1} + a_{m+2} + a_{m+3} + \dots) = 0.$$

В заключение докажем одно, часто используемое, необходимое условие сходимости ряда.

Теорема 6. *Общий член сходящегося ряда при возрастании своего номера стремится к нулю.*

Иными словами, из сходимости ряда

$$a_1 + a_2 + a_3 + \dots \tag{22}$$

вытекает, что

$$\lim_{n\to\infty} a_n = 0. \tag{23}$$

Действительно, если сумму ряда (22) обозначить через S, то с возрастанием n каждая из сумм

$$S_n = a_1 + a_2 + \dots + a_n, \quad S_{n-1} = a_1 + a_2 + \dots + a_{n-1}$$

будет стремиться к S. Но тогда

$$S_n - S_{n-1} \to S - S = 0,$$

и остаётся заметить, что $S_n - S_{n-1} = a_n$.

Весьма важно подчеркнуть, что условие (23), будучи необходимым для сходимости ряда (22), вовсе не является для этой сходимости достаточным. В самом деле, гармонический ряд

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots$$

очевидно удовлетворяет условию (23), но, как было показано в п° 33, он расходится.

35. Положительные ряды. Если

$$a_n \geqslant 0 \qquad (n = 1, 2, 3, \dots),$$

то ряд $a_1 + a_2 + a_3 + \dots$ называется *положительным*. В том случае, когда при всех n оказывается $a_n > 0$, мы будем называть ряд *строго положительным*.

Положительные ряды обладают многими свойствами, сближающими их с обычными суммами конечного числа слагаемых.

Легко видеть, что частичная сумма

$$S_n = a_1 + a_2 + \ldots + a_n$$

положительного ряда возрастает (может быть, не строго) с увеличением n. Так как всякая возрастающая числовая последовательность имеет конечный или бесконечный предел (причём члены последовательности не превосходят этого предела), то для любого положительного ряда существует предел

$$S = \lim_{n\to\infty} S_n.$$

Этот предел будет конечным или бесконечным, смотря по тому, ограничено сверху или нет множество частичных сумм $\{S_n\}$. Таким образом, имеет место

Теорема 1. *Положительный ряд сходится тогда и только тогда, когда множество его частичных сумм ограничено сверху.*

Разумеется, у ряда не положительного ограниченность множества частичных сумм не обеспечивает сходимости, как это видно из примера ряда $1+(-1)+1+(-1)+\ldots$

Отметим ещё, что частичные суммы сходящегося положительного ряда не превосходят его суммы.

Доказанная теорема сводит вопрос о сходимости положительного ряда к более простому вопросу об ограниченности множества его частичных сумм.

Рассмотрим, например, ряд

$$1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\frac{1}{4^\alpha}+\ldots, \tag{24}$$

в котором $\alpha > 1$. Сумму S_{2^m-1} этого ряда можно записать так:

$$S_{2^m-1} = 1+\left[\frac{1}{2^\alpha}+\frac{1}{3^\alpha}\right]+\left[\frac{1}{4^\alpha}+\frac{1}{5^\alpha}+\frac{1}{6^\alpha}+\frac{1}{7^\alpha}\right]+ \\ +\ldots+\left[\frac{1}{(2^{m-1})^\alpha}+\ldots+\frac{1}{(2^m-1)^\alpha}\right].$$

Так как сумма

$$\frac{1}{(2^k)^\alpha}+\frac{1}{(2^k+1)^\alpha}+\ldots+\frac{1}{(2^{k+1}-1)^\alpha}$$

содержит 2^k слагаемых, а самое большое из них есть первое, то эта сумма не превосходит числа

$$\frac{1}{(2^k)^\alpha}\cdot 2^k = \frac{1}{(2^{\alpha-1})^k}.$$

Поэтому

$$S_{2^m-1} < 1+\frac{1}{2^{\alpha-1}}+\frac{1}{(2^{\alpha-1})^2}+\ldots+\frac{1}{(2^{\alpha-1})^{m-1}}.$$

Стоящая здесь справа сумма есть частичная сумма геометрической прогрессии

$$1+\frac{1}{2^{\alpha-1}}+\frac{1}{(2^{\alpha-1})^2}+\frac{1}{(2^{\alpha-1})^3}+\ldots \tag{25}$$

В силу сказанного в конце п° 33 эта прогрессия сходится (ибо $\alpha>1$), и сумма её равна

$$\frac{2^{\alpha-1}}{2^{\alpha-1}-1}. \tag{26}$$

Так как прогрессия (25) также является рядом положительным, то её частичные суммы не превосходят её суммы (26). Тем более

$$S_{2^m-1}<\frac{2^{\alpha-1}}{2^{\alpha-1}-1}.$$

Это неравенство установлено для любого m. Но для всякого n можно найти такое m, что

$$2^m-1>n.$$

Поэтому при всяком n оказывается

$$S_n<\frac{2^{\alpha-1}}{2^{\alpha-1}-1},$$

и ряд (24) сходится.

Следует, однако, заметить, что непосредственное применение теоремы 1 встречается сравнительно редко.

Обычно применяют основанные на ней, но более удобные признаки сходимости рядов. Простейший из них — это так называемый признак сравнения рядов.

Определение. Если каждый член положительного ряда не больше, чем имеющий тот же номер член другого ряда, то второй ряд называется *мажорантным* по отношению к первому.

Иначе говоря, ряд

$$b_1+b_2+b_3+\ldots$$

является мажорантным по отношению к ряду

$$a_1+a_2+a_3+\ldots,$$

если при всех n будет

$$a_n\leqslant b_n.$$

Легко понять, что частичная сумма данного ряда не больше, чем (имеющая тот же номер) частичная сумма ряда мажорантного. Значит, если ограничены сверху частичные суммы мажорантного ряда, то это и подавно так для исходного ряда. Отсюда вытекает

Теорема 2. *Если для положительного ряда существует сходящийся мажорантный ряд, то и сам этот ряд сходится.*

Если же данный ряд расходится, то расходится и всякий мажорантный для него ряд.

Рассмотрим, например, ряд

$$1+\frac{1}{2^{\alpha}}+\frac{1}{3^{\alpha}}+\frac{1}{4^{\alpha}}+\ldots, \tag{27}$$

предполагая $\alpha<1$. Ясно, что этот ряд — мажорантный по отношению к гармоническому ряду, и потому ряд (27) расходится.

Исследование сходимости какого-либо ряда при помощи теоремы 2 наталкивается на необходимость нахождения другого ряда, поведение которого в отношении его сходимости нам известно и с которым мы хотим сравнить наш исходный ряд. Нахождение такого «ряда сравнения» в значительной степени зависит от проницательности исследователя, и в этом существенный недостаток признака сравнения. Существуют признаки сходимости, носящие гораздо более алгорифмический характер. Мы остановимся только на одном из таких признаков — на признаке Даламбера.

Лемма 1. *Если строго положительный ряд*

$$a_1+a_2+a_3+\ldots \tag{28}$$

таков, что при всех n оказывается

$$\frac{a_{n+1}}{a_n}>1,$$

то этот ряд расходится.

В самом деле, здесь $a_{n+1}>a_n$, т. е. общий член ряда возрастает с увеличением своего номера и не может стремиться к нулю. Иначе говоря, не выполнено необходимое условие сходимости ряда.

Лемма 2. *Если строго положительный ряд* (28) *таков, что при всех n оказывается*

$$\frac{a_{n+1}}{a_n}\leqslant q,$$

где q — некоторое постоянное число, меньшее единицы

$$0<q<1,$$

то этот ряд сходится.

Действительно, по условию

$$\frac{a_2}{a_1}\leqslant q,\quad \frac{a_3}{a_2}\leqslant q,\quad \ldots,\quad \frac{a_n}{a_{n-1}}\leqslant q.$$

Перемножая эти неравенства, после надлежащих сокращений получим:

$$\frac{a_n}{a_1}\leqslant q^{n-1},$$

откуда

$$a_n\leqslant a_1 q^{n-1}.$$

Последнее неравенство показывает, что члены ряда (28) не больше соответствующих членов ряда

$$a_1 + a_1 q + a_1 q^2 + a_1 q^3 + \dots \tag{29}$$

Но ряд (29) есть геометрическая прогрессия. Так как знаменатель этой прогрессии (по абсолютной величине) меньше 1, то она сходится. Но тогда и подавно сходится ряд (28).

Читатель обратит внимание на то, что неравенство

$$\frac{a_{n+1}}{a_n} < 1$$

ещё не обеспечивает сходимости ряда (28). Это видно хотя бы на примере гармонического ряда.

Теорема 3. (Признак Даламбера.) *Допустим, что строго положительный ряд* (29) *таков, что существует* (*конечный или бесконечный*) *предел*

$$l = \lim_{n \to \infty} \frac{a_{n+1}}{a_n} \tag{30}$$

Тогда при $l > 1$ *ряд* (28) *расходится, а при* $l < 1$ *сходится.*

Допустим сначала, что $l > 1$. Так как дробь $\frac{a_{n+1}}{a_n}$ стремится к l, то достаточно далёкие значения этой дроби будут удовлетворять неравенству

$$\frac{a_{n+1}}{a_n} > 1. \tag{31}$$

Пусть, например, это неравенство выполнено для всех n, удовлетворяющих неравенству $n > m$. Тогда ряд

$$a_{m+1} + a_{m+2} + a_{m+3} + \dots \tag{32}$$

таков, что у него отношение уже *любого* последующего члена к своему предыдущему оказывается удовлетворяющим неравенству (31). Значит (по лемме 1), ряд (32) расходится, а так как это есть остаток ряда (28), то этот последний ряд также расходится.

Пусть теперь $l < 1$. Закрепим какое-нибудь число q, удовлетворяющее неравенству $l < q < 1$ (например, положим $q = \frac{l+1}{2}$). Тогда найдётся такое m, что при всех $n > m$ будет

$$\frac{a_{n+1}}{a_n} < q.$$

Снова составляя ряд (32) и применяя к нему лемму 2, убеждаемся сначала в сходимости ряда (32), а затем и ряда (28).

Доказанная теорема действительно имеет совершенно алгорифмический характер: для её применения надо лишь составить отно-

шение $\frac{a_{n+1}}{a_n}$ и изучить его поведение при безгранично возрастающем n. Никаких вспомогательных рядов для сопоставления с данным рядом искать уже не требуется. Надо заметить, однако, что теорема 3 применима далеко не всегда. Не говоря уже о том, что предела (30) может не существовать, этот предел может равняться 1, и тогда теорема также не позволяет сделать никакого заключения относительно сходимости ряда.

Рассмотрим несколько примеров.

Пример 1. Рассмотрим ряд

$$\frac{1!}{1}+\frac{2!}{2^2}+\frac{3!}{3^3}+\frac{4!}{4^4}+\dots .$$

Здесь

$$a_n=\frac{n!}{n^n},\quad a_{n+1}=\frac{(n+1)!}{(n+1)^{n+1}},\quad \frac{a_{n+1}}{a_n}=\left(\frac{n}{n+1}\right)^n=\frac{1}{\left(1+\frac{1}{n}\right)^n}.$$

Поэтому

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\frac{1}{e}=\frac{1}{2{,}71828\dots}<1,$$

и ряд сходится.

Пример 2. Для ряда

$$1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\frac{1}{4^\alpha}+\dots \tag{33}$$

будет

$$a_n=\frac{1}{n^\alpha},\quad a_{n+1}=\frac{1}{(n+1)^\alpha},\quad \frac{a_{n+1}}{a_n}=\left(\frac{n}{n+1}\right)^\alpha\to 1.$$

Поэтому признак Даламбера не даёт возможности ответить на вопрос о сходимости ряда (33). Мы уже знаем, что этот ряд сходится при $\alpha>1$ и расходится при $\alpha\leqslant 1$. Таким образом, бывают и сходящиеся и расходящиеся ряды, у которых $l=1$.

Пример 3. Исследовать, для каких $x\,(x>0)$ сходится ряд

$$\frac{x}{1^2}+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\frac{x^4}{4^2}+\dots$$

Здесь

$$a_n=\frac{x^n}{n^2},\quad a_{n+1}=\frac{x^{n+1}}{(n+1)^2},\quad \frac{a_{n+1}}{a_n}=\left(\frac{n}{n+1}\right)^2 x\to x.$$

Следовательно, ряд сходится для $x<1$ и расходится для $x>1$. При $x=1$ признак Даламбера ответа не даёт, но непосредственно ясно, что ряд сходится.

Помимо признака Даламбера, являющегося простейшим, существует большое число и других признаков сходимости положительных рядов: признак Коши, признак Раабе, интегральный признак и др. Очень

общий признак был предложен известным русским математиком, профессором Киевского университета В. П. Ермаковым (1845—1922).

В заключение докажем одно простое, но важное свойство положительных рядов, которое будет использовано ниже.

Теорема 4. *Если в сходящемся положительном ряде вычеркнуть любое множество членов и составить ряд из оставшихся членов, оставляя их в том же порядке, что и в исходном ряде, то вновь полученный ряд также будет сходиться.*

Действительно, пусть положительный ряд

$$a_1 + a_2 + a_3 + a_4 + \dots \tag{34}$$

сходится и имеет сумму S.

Рассмотрим последовательность натуральных чисел

$$n_1 < n_2 < n_3 < \dots$$

и образуем ряд

$$a_{n_1} + a_{n_2} + a_{n_3} + \dots \tag{35}$$

Теорема утверждает, что ряд (35) сходится. Чтобы доказать это, обозначим через S_n и S_k^* частичные суммы рядов (34) и (35):

$$S_n = a_1 + a_2 + \dots + a_n, \quad S_k^* = a_{n_1} + a_{n_2} + \dots + a_{n_k}.$$

Легко видеть, что

$$S_k^* \leqslant S_{n_k}.$$

Отсюда следует, что $S_k^* \leqslant S$, а потому ряд (35) сходится.

36. Знакочередующиеся ряды. Переходя к рассмотрению рядов, члены которых уже не обязательно положительны, остановимся сначала на одном важном частном типе этих рядов — на рядах знакочередующихся, теория которых сравнительно проста.

Определение. Ряд называется *знакочередующимся*, если любые два его соседних члена суть числа разных знаков.

Несколько изменяя употреблявшуюся выше символику, будем теперь обозначать через a_n не сам общий член ряда, а его абсолютную величину. Тогда, предполагая для определённости, что первый член знакочередующегося ряда положителен, мы сможем записать этот ряд в форме [1])

$$a_1 - a_2 + a_3 - a_4 + a_5 - \dots \tag{36}$$

Теорема Лейбница. *Если абсолютная величина общего члена знакочередующегося ряда убывает и стремится к нулю, то этот ряд сходится.*

---

[1]) Строго говоря, введение этой записи представляет новое соглашение, ибо согласно определению ряда мы должны вместо (36) писать

$$a_1 + (-a_2) + a_3 + (-a_4) + \dots$$

Действительно, допустим, что ряд (36) таков, что

$$a_1 > a_2 > a_3 > a_4 > \dots, \tag{37}$$

$$\lim_{n\to\infty} a_n = 0. \tag{38}$$

Образуем частичные суммы S_{2n}:

$$S_2 = (a_1 - a_2),$$
$$S_4 = (a_1 - a_2) + (a_3 - a_4),$$
$$S_6 = (a_1 - a_2) + (a_3 - a_4) + (a_5 - a_6),$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$

Благодаря (37), все скобки положительны. Значит,

$$S_2 < S_4 < S_6 < \dots$$

Иначе говоря, последовательность $\{S_{2n}\}$ возрастает. С другой стороны,

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \dots - (a_{2n-2} - a_{2n-1}) - a_{2n},$$

откуда ясно, что

$$S_{2n} < a_1.$$

Как известно, при этих условиях существует конечный предел

$$S = \lim_{n\to\infty} S_{2n}.$$

Но

$$S_{2n+1} = S_{2n} + a_{2n+1},$$

откуда в связи с (38) вытекает, что сумма S_{2n+1} с возрастанием n также стремится к S. Итак, при достаточно больших n сумма S_n будет сколь угодно близка к S независимо от чётности n. Иначе говоря,

$$\lim_{n\to\infty} S_n = S,$$

чем и доказана теорема.

Заметим, что теорема перестаёт быть верной, если отбросить условие убывания a_n. Например, знакочередующийся ряд

$$1 - \frac{1}{10} + \frac{1}{2} - \frac{1}{100} + \frac{1}{3} - \frac{1}{1000} + \dots,$$

как легко видеть, расходится [1]).

---

[1]) В самом деле, у этого ряда

$$S_{2n} = \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}\right) - \left(\frac{1}{10} + \frac{1}{10^2} + \dots + \frac{1}{10^n}\right).$$

Первая скобка неограниченно растёт, а вторая не больше, чем $\frac{1}{9}$.

Из доказательства теоремы видно, что сумма S_{2n} стремится к своему пределу S возрастая. Таким образом, S_{2n} есть значение S по недостатку. Напротив, S_{2n+1} есть значение S по избытку. Действительно, из равенств

$$\begin{aligned} S_1 &= a_1, \\ S_3 &= a_1 - (a_2 - a_3), \\ S_5 &= a_1 - (a_2 - a_3) - (a_4 - a_5), \\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

видно, что $S_1 > S_3 > S_5 > \ldots$, так что сумма S_{2n+1} стремится к S убывая.

Последние два свойства знакочередующегося ряда, удовлетворяющего условиям теоремы Лейбница, можно представить и в более наглядной форме. Именно, из того, что $S_2 < S < S_1$, следует, что

$$0 < S < a_1.$$

Значит, сумма нашего ряда имеет знак его первого члена, а по абсолютной величине меньше этого члена. Но ведь остаток ряда (36) после n-го члена

$$(-1)^n [a_{n+1} - a_{n+2} + a_{n+3} - \ldots]$$

сам тоже есть ряд такого же типа. Поэтому сумма R_n этого остатка имеет знак $(-1)^n$ и $|R_n| < a_{n+1}$. Если вспомнить, что

$$S = S_n + R_n,$$

то всё сказанное можно будет резюмировать так: *отбрасывая все члены ряда* (36) *после* n-го *и принимая частичную сумму*

$$S_n = a_1 - a_2 + a_3 - \ldots \pm a_n$$

за сумму S, *мы совершаем ошибку* [1]), *которая имеет тот же знак, что и первый отброшенный член* $(-1)^n a_{n+1}$, *а по абсолютной величине меньше его.*

Это свойство рядов, удовлетворяющих условиям теоремы Лейбница, позволяет уверенно применять такие ряды в приближённых подсчётах.

Пусть, например,

$$S = 1 - \frac{1}{2^2} + \frac{1}{3^3} - \frac{1}{4^4} + \frac{1}{5^5} - \ldots$$

Ограничиваясь выписанными членами для нахождения суммы этого ряда, мы сделаем отрицательную ошибку, по абсолютной величине меньшую, чем

$$\frac{1}{6^6} = \frac{1}{46\,656} < 0{,}000025.$$

---

[1]) Мы называем ошибкой число, которое надо прибавить к S_n, чтобы получить S (т. е. R_n).

37. Абсолютная сходимость. Теперь рассмотрим такие ряды, знаки членов которых уже совершенно произвольны. При этом мы снова будем обозначать через a_1, a_2, a_3, ... сами члены ряда.

Теорема 1. *Сопоставим с рядом*

$$a_1 + a_2 + a_3 + \dots \tag{39}$$

ряд

$$|a_1| + |a_2| + |a_3| + \dots, \tag{40}$$

составленный из абсолютных величин членов данного ряда. Если сходится ряд (40), *то сходится и исходный ряд* (39).

В самом деле, пусть ряд

$$b_1 + b_2 + b_3 + \dots \tag{41}$$

есть ряд, состоящий из всех положительных (или равных нулю) членов нашего ряда (39) [причём их взаимное расположение таково же, как и в ряде (39)]. Пусть, далее,

$$c_1 + c_2 + c_3 + \dots \tag{42}$$

есть ряд [1]) абсолютных величин отрицательных членов ряда (39) (также расположенных в том порядке, в котором эти члены следуют друг за другом в исходном ряде).

Каждый из рядов (41) и (42) получается из сходящегося положительного ряда (40) путём вычёркивания части его членов (например, чтобы из (40) получить (41), надо вычеркнуть из (40) числа c_1, c_2, c_3, ...). Поэтому в силу теоремы 4 п° 35 ряды (41) и (42) сходятся. Обозначим их суммы соответственно через B и C.

Обозначим, далее, через A_n, B_n и C_n частичные суммы рядов (39), (41) и (42). Пусть среди чисел

$$a_1, a_2, a_3, \dots, a_n$$

имеется $m(n)$ неотрицательных и $p(n)$ отрицательных.

Тогда

$$A_n = B_{m(n)} - C_{p(n)}.$$

Правая часть этого равенства с ростом n стремится к разности $B - C$. Значит, и левая часть стремится к тому же пределу. Теорема доказана.

Заметим, что из сходимости ряда (39) не вытекает, что сходится (40). Например, ряд

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

---

[1]) Мы считаем, что в (39) имеется бесконечное множество положительных и бесконечное множество отрицательных членов, ибо иначе всё становится тривиальным.

сходится (это следует хотя бы из теоремы Лейбница), но ряд, составленный из абсолютных величин, будучи гармоническим, расходится.

Таким образом, требование сходимости ряда (40) представляет собой более тяжёлое требование, чем требование сходимости ряда (39). В связи с этим такой ряд (39), который не только сходится сам, но для которого сходится и ряд абсолютных величин, называется *абсолютно сходящимся*. Если же ряд (39) сходится, но ряд (40) расходится, то говорят, что (39) есть ряд неабсолютно сходящийся.

Из приведённого доказательства теоремы 1 вытекает справедливость и такого предложения:

Теорема 2. *Если ряд* (39) *сходится абсолютно, то сходятся также ряды* (41) *и* (42), *образованные положительными членами и модулями отрицательных членов этого ряда. Суммы всех трёх рядов связаны соотношением*

$$A = B - C. \tag{43}$$

Признак сходимости Даламбера переносится в теорию рядов с членами любых знаков в следующей форме:

Теорема 3. *Пусть ряд* (39) *таков, что существует предел*

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = l.$$

Если $l < 1$, *то ряд сходится, а если* $l > 1$, *то расходится.*

Действительно, если $l < 1$, то по признаку Даламбера будет сходиться ряд абсолютных величин членов данного ряда, а значит, и подавно и сам ряд (39). Если же $l > 1$, то найдётся такое m, что при $n \geqslant m$ будет

$$\left|\frac{a_{n+1}}{a_n}\right| > 1.$$

Но тогда

$$|a_m| < |a_{m+1}| < |a_{m+2}| < \dots,$$

и общий член ряда (39) не стремится к нулю, откуда вытекает расходимость этого ряда.

Примеры применения этой теоремы будут приведены ниже.

38. Вопрос о перестановке членов ряда. Умножение рядов. Мы уже говорили выше, что с чисто формальной точки зрения ряды — это не что иное, как суммы, содержащие бесконечное множество слагаемых. Далее отмечалось, что этой формальной точке зрения не следует придавать слишком большого значения. Хорошей иллюстрацией к этому замечанию может служить вопрос о перестановке членов ряда.

Одним из самых основных законов арифметики является переместительный закон: «сумма не зависит от порядка слагаемых». Оказывается, что для бесконечных рядов этот закон уже неверен! Покажем это на следующем примере.

Согласно теореме Лейбница знакочередующийся ряд

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\dots \tag{44}$$

сходится. Обозначим сумму его через S. Важно отметить, что $S\neq 0$, ибо сумма ряда, удовлетворяющего условиям теоремы Лейбница, имеет знак первого члена ряда.

Образуем теперь ряд из тех же чисел, но располагая их так, чтобы за каждым положительным членом следовало два отрицательных

$$1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\frac{1}{5}-\frac{1}{10}-\frac{1}{12}+\dots \tag{45}$$

Обозначим через S_n и S_n^* частичные суммы рядов (44) и (45). Тогда, объединяя члены в группы по три, получим [1]):

$$S_{3n}^*=\left(1-\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{6}-\frac{1}{8}\right)+\dots+\left(\frac{1}{2n-1}-\frac{1}{4n-2}-\frac{1}{4n}\right).$$

Замечая, что

$$1-\frac{1}{2}=\frac{1}{2},\quad \frac{1}{3}-\frac{1}{6}=\frac{1}{6},\quad \dots,\quad \frac{1}{2n-1}-\frac{1}{4n-2}=\frac{1}{4n-2},$$

находим:

$$S_{3n}^*=\left(\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{6}-\frac{1}{8}\right)+\dots+\left(\frac{1}{4n-2}-\frac{1}{4n}\right),$$

откуда

$$S_{3n}^*=\frac{1}{2}\left[1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\dots+\frac{1}{2n-1}-\frac{1}{2n}\right].$$

Но это означает, что

$$S_{3n}^*=\frac{1}{2}S_{2n},$$

и потому S_{3n}^* с возрастанием n стремится к $\frac{1}{2}S$.

Так как

$$S_{3n+1}^*=S_{3n}^*+\frac{1}{2n+1},\quad S_{3n+2}^*=S_{3n+1}^*-\frac{1}{4n+2},$$

то и эти суммы стремятся к $\frac{1}{2}S$. Итак, ряд (45) имеет сумму $\frac{1}{2}S$, и эта сумма **не равна** сумме ряда (44) (ибо $S\neq 0$). Мы видим, что перестановка членов сходящегося ряда может изменить его сумму.

Исследуем затронутый вопрос, ввиду его большого методологического интереса, подробнее. Оказывается, что для **положительных** рядов переместительный закон сохраняется.

Теорема 1. *От перестановки членов сходящегося положительного ряда сходимость его не нарушается, а сумма не меняется.*

---

[1]) Состав k-й скобки таков: $\frac{1}{2k-1}-\frac{1}{4k-2}-\frac{1}{4k}$. Действительно, k-е нечётное натуральное число есть $2k-1$, а $2k$-е чётное есть $4k$.

В самом деле, пусть положительный ряд

$$a_1 + a_2 + a_3 + \dots \tag{46}$$

сходится. Обозначим его сумму через S. Образуем новый ряд

$$a_{n_1} + a_{n_2} + a_{n_3} + \dots, \tag{47}$$

состоящий из тех же членов, что и (44), но расположенных в другом порядке. Пусть частичные суммы рядов (46) и (47) суть S_n и $S_{n'}^*$. Пусть $m(k)$ есть наибольшее из чисел $n_1, n_2, \dots, n_k$. Тогда все слагаемые суммы

$$S_k^* = a_{n_1} + a_{n_2} + \dots + a_{n_k}$$

будут входить в сумму

$$S_{m(k)} = a_1 + a_2 + \dots + a_{m(k)},$$

где, однако, могут находиться ещё и такие (положительные) слагаемые, которых нет в сумме S_k^*. Отсюда следует, что

$$S_k^* \leqslant S_{m(k)}$$

и тем более

$$S_k^* \leqslant S. \tag{48}$$

Таким образом, множество частичных сумм положительного ряда (47) ограничено сверху, и этот ряд сходится. Значит, мы доказали, что сходимость положительного ряда от перестановки его членов не нарушается. Чтобы установить неизменность суммы ряда, обозначим сумму ряда (47) через S^*. Тогда, переходя к пределу в неравенстве (48), мы найдём

$$S^* \leqslant S.$$

Это значит, что сумма сходящегося положительного ряда при перестановке его членов не увеличивается. Но тогда уже ясно, что она и не уменьшается, ибо иначе обратная перестановка приводила бы к увеличению суммы ряда. Теорема доказана.

Эта теорема обобщается на любые абсолютно сходящиеся ряды.

Теорема 2. *От перестановки членов абсолютно сходящегося ряда его абсолютная сходимость не нарушается и сумма не меняется.*

Абсолютная сходимость ряда

$$a_1 + a_2 + a_3 + \dots \tag{49}$$

означает сходимость положительного ряда

$$|a_1| + |a_2| + |a_3| + \dots \tag{50}$$

Пусть в результате перестановки своих членов ряд (49) превратился в ряд

$$a_{n_1} + a_{n_2} + a_{n_3} + \dots \tag{51}$$

Ясно, что ряд абсолютных величин членов ряда (51) получается некоторой перестановкой из ряда (50). Значит, этот ряд абсолютных величин сходится (по предыдущей теореме), а это и означает, что ряд (51) абсолютно сходится.

Введём, далее, ряды

$$b_1 + b_2 + b_3 + \dots, \tag{52}$$

$$c_1 + c_2 + c_3 + \dots, \tag{53}$$

состоящие из неотрицательных и модулей отрицательных членов ряда (49). Эти ряды сходятся, и их суммы B и C связаны с суммой A ряда (49) формулой

$$A = B - C. \tag{54}$$

Когда мы производим перестановку членов в ряде (49), то это вызывает соответствующие перестановки и в рядах (52) и (53), ибо в этих рядах порядок членов такой же, как и в (49). Но так как ряды (52) и (53) — положительные, то их суммы B и C от перестановки не меняются. С другой стороны, сумма ряда (51) должна выражаться через суммы рядов, полученных упомянутыми перестановками из (52) и (53), той же формулой (54). Отсюда и видно, что сумма ряда (51) равна сумме ряда (49).

Условие абсолютной сходимости ряда для законности перестановок его членов не только достаточно, но и необходимо. Действительно, имеет место

Теорема 3. *Если ряд* (49) *сходится неабсолютно, то его члены можно переставить так, чтобы вновь полученный ряд имел любую наперёд заданную сумму, а также так, чтобы он расходился.*

Мы не будем останавливаться на доказательстве этого предложения.

Обратимся теперь к вопросу о перемножении рядов.

Для того, чтобы перемножить две конечные суммы

$$A = a_1 + a_2 + \ldots + a_n, \quad B = b_1 + b_2 + \ldots + b_m,$$

надо умножить каждое слагаемое первой суммы на каждое слагаемое второй суммы и составить сумму всевозможных произведений

$$a_i b_k \qquad (i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m). \tag{55}$$

Естественно спросить, переносится ли это правило и в теорию бесконечных рядов? Здесь прежде всего мы сталкиваемся с тем обстоятельством, что появляется бесконечное множество произведений (55), и потому нам приходится дать какой-либо способ их «сложения». Наиболее удобным является следующее определение:

Произведением рядов

$$a_1 + a_2 + a_3 + \ldots \tag{56}$$

и

$$b_1 + b_2 + b_3 + \ldots \tag{57}$$

называется ряд

$$c_1 + c_2 + c_3 + \ldots, \tag{58}$$

в котором

$$c_1 = a_1 b_1, \quad c_2 = a_1 b_2 + a_2 b_1, \quad c_3 = a_1 b_3 + a_2 b_2 + a_3 b_1,$$

и вообще

$$c_n = a_1 b_n + a_2 b_{n-1} + \ldots + a_n b_1. \tag{59}$$

Оказывается, что, вообще говоря, из сходимости рядов (56) и (57) не вытекает сходимость ряда (58). Например, если умножить сам на себя сходящийся [1]) ряд

$$1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \ldots,$$

то абсолютная величина общего члена (59) ряда (58) будет такова

$$|c_n| = 1 \cdot \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{n-2}} + \ldots + \frac{1}{\sqrt{n}} \cdot 1.$$

[1]) Это видно из теоремы Лейбница, доказанной в п° 36.

Отсюда видно, что

$$|c_n| \geqslant \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{n}} + \ldots + \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{n}} = 1,$$

и ряд (58) расходится.

Тем не менее справедлива

Теорема 4. *Если ряды* (56) *и* (57) *сходятся абсолютно и их суммы соответственно равны A и B, то и ряд* (58) *сходится абсолютно и сумма его C равна произведению сумм рядов* (56) *и* (57)

$$C = AB. \tag{60}$$

Иными словами, абсолютно сходящиеся ряды можно перемножать, как конечные суммы.

Для доказательства этой теоремы расположим все произведения (55) в бесконечную матрицу

$$\begin{array}{llll} a_1 b_1 & a_1 b_2 & a_1 b_3 \ldots & a_1 b_m \ldots \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \ldots & a_2 b_m \ldots \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \ldots & a_3 b_m \ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_n b_1 & a_n b_2 & a_n b_3 \ldots & a_n b_m \ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \end{array} \tag{61}$$

Элементы этой матрицы можно различными способами записать в форме последовательности. Для нас особенно важны способы, изображённые на схемах *I* и *II*

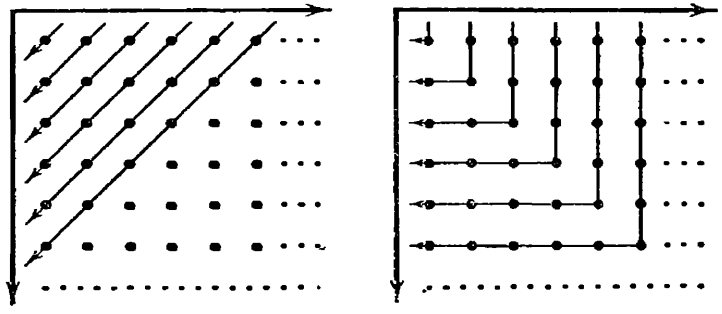


Схема I *Схема II*

Первый из этих способов приводит к ряду

$$a_1 b_1 + a_1 b_2 + a_2 b_1 + a_1 b_3 + a_2 b_2 + a_3 b_1 + a_1 b_4 + a_2 b_5 + \ldots \tag{62}$$

Объединяя члены ряда (62) в группы из одного, двух, трёх и т. д. членов, мы приходим в точности к ряду (58).

Способ *II* приводит к ряду

$$a_1 b_1 + a_1 b_2 + a_2 b_2 + a_2 b_1 + a_1 b_3 + a_2 b_3 + a_3 b_3 + a_3 b_2 + \ldots \tag{63}$$

Все ряды, получаемые расположением элементов матрицы (61) в форме последовательности, получаются друг из друга перестановками членов. Покажем, что все эти ряды сходятся абсолютно. Именно, возьмём какой-нибудь

из этих рядов, составим ряд абсолютных величин членов взятого ряда и обозначим через S_n^* частичную сумму этого ряда абсолютных величин. Если n закреплено, а m достаточно велико, то все слагаемые суммы S_n^* будут содержаться среди чисел, заполняющих квадрат

$$\begin{array}{llll} |a_1 b_1| & |a_1 b_2| & \dots & |a_1 b_m| \\ |a_2 b_1| & |a_2 b_2| & \dots & |a_2 b_m| \\ \dots\dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ |a_m b_1| & |a_m b_2| & \dots & |a_m b_m| \end{array}$$

Поэтому сумма S_n^* будет не больше, чем сумма в с е х чисел, входящих в указанный квадрат, т. е. не больше, чем

$$(|a_1|+|a_2|+\dots+|a_m|)\,(|b_1|+|b_2|+\dots+|b_m|).$$

Если обозначить через A^* и B^* суммы рядов, составленных из абсолютных величин членов рядов (56) и (57), то из сказанного ясно, что

$$S_n^* \leqslant A^* B^*.$$

Отсюда и следует наше утверждение об абсолютной сходимости рядов, порождаемых матрицей (61).

В частности, абсолютно сходится и ряд (62). Это означает, что сходится ряд

$$|a_1 b_1|+|a_1 b_2|+|a_2 b_1|+|a_1 b_3|+|a_2 b_2|+|a_3 b_1|+|a_1 b_4|+\dots \qquad (64)$$

Объединяя члены ряда (64) в группы из одного, двух, трёх и т. д. членов, мы убеждаемся в сходимости ряда

$$d_1+d_2+d_3+\dots,$$

где

$$d_n=|a_1 b_n|+|a_2 b_{n-1}|+\dots+|a_n b_1|.$$

Но ведь $|c_n| \leqslant d_n$. Значит, ряд (58) и подавно абсолютно сходится. Остаётся доказать равенство (60).

Так как все ряды, получаемые расположением членов матрицы (61) в какую-нибудь последовательность, сходятся абсолютно, а с другой стороны, они получаются и друг из друга перестановками членов, то все они имеют одну и ту же сумму S. В частности, эту сумму S имеют и ряды (62) и (63). Но ряд (58) получается объединением членов ряда (62) в надлежащие группы, что, как мы знаем, не меняет суммы сходящегося ряда. Значит, $S=C$. Таким образом, C есть сумма ряда (63). Эту последнюю сумму мы найдём, объединив члены ряда (63) в группы из одного, трёх, пяти и т. д. членов, т. е. как сумму ряда

$$a_1 b_1+(a_1 b_2+a_2 b_2+a_2 b_1)+(a_1 b_3+a_2 b_3+a_3 b_3+a_3 b_2+a_3 b_1)+\dots$$

Сумма n членов этого последнего ряда, очевидно, может быть записана так:

$$\begin{array}{c} a_1 b_1+a_1 b_2+\dots+a_1 b_n+ \\ +a_2 b_1+a_2 b_2+\dots+a_2 b_n+ \\ \dots\dots\dots\dots\dots\dots \\ +a_n b_1+a_n b_2+\dots+a_n b_n, \end{array}$$

откуда ясно, что она равна произведению

$$(a_1+a_2+\dots+a_n)\,(b_1+b_2+\dots+b_n)$$

частичных сумм рядов (56) и (57). Увеличивая n и переходя к пределу, мы и приходим к (60).

Приведём без доказательства ещё два предложения:

Теорема 5. *Если ряды* (56) *и* (57) *сходятся и хоть один из них сходится абсолютно, то сходится* (хотя и не обязательно абсолютно) *и ряд* (58), *и суммы рядов* (56), (57) *и* (58) *связаны равенством* (60).

Теорема 6. *Если сходятся все три ряда* (56), (57) *и* (58), *то справедливо* (60).

§ 8. Степенные ряды

39. Промежуток сходимости. *Степенным рядом* называется ряд вида

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots, \qquad (1)$$

где c_0, c_1, c_2, ... — постоянные числа, называемые коэффициентами ряда. Про ряд (1) говорят, что он расположен по степеням x. Иногда рассматривают степенные ряды несколько более общего вида

$$c_0 + c_1 (x-a) + c_2 (x-a)^2 + \dots, \qquad (2)$$

расположенные по степеням разности $x-a$. Впрочем, ряд (2) приводится к виду (1) при помощи подстановки $x-a=x'$. Ввиду этого обстоятельства мы занимаемся ниже преимущественно рядами (1).

Мы говорили, что с чисто формальной точки зрения бесконечные ряды представляют собой суммы, состоящие из бесконечного числа слагаемых. Аналогично этому, степенные ряды — это, так сказать, «многочлены бесконечно высокой степени». Важность таких выражений для математики видна хотя бы из следующего примера: в п° 12 мы установили, что при любом действительном x будет

$$\sin x = \lim_{p\to\infty}\left[x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^p \frac{x^{2p+1}}{(2p+1)!}\right].$$

Пользуясь обозначениями настоящей главы, мы можем этот результат записать и так:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \qquad (3)$$

Иначе говоря, важнейшая тригонометрическая функция $\sin x$ представима степенным рядом (1).

Точно так же из рассуждений п°12 вытекает, что

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots, \qquad (4)$$

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \qquad (5)$$

Мы видим, что степенные ряды представляют собой выражения, способные изображать многие важные функции.

Изображение функции в виде суммы степенного ряда может быть полезно в разных отношениях. Прежде всего оно даёт возможность фактического вычисления этой функции, ибо частичные суммы ряда, будучи обыкновенными алгебраическими многочленами, вычисляются без труда, и в то же время эти суммы становятся сколь угодно близкими к изображаемой функции при достаточном увеличении числа взятых членов. Кроме того, представление функции степенным рядом часто позволяет устанавливать различные свойства этой функции. Так, например, из формул (3) и (4) получаем:

$$\sin(-x) = -\sin x, \quad \cos(-x) = \cos x,$$

т. е. что $\sin x$ есть функция нечётная, а $\cos x$ — чётная. Мы надеемся, что изложенные здесь соображения достаточно объясняют тот большой интерес, который привлекает к себе теория степенных рядов. Ниже важность этой теории (и, в частности, её важность для нужд элементарной математики) будет выявлена более обстоятельно.

По отношению к степенному ряду (как и вообще ко всякому ряду, члены которого зависят от x) бессмысленно ставить вопрос, сходится он или расходится, ибо один и тот же ряд при одних значениях x может сходиться, а при других — расходиться. Например, ряд

$$1 + 2x + 3x^2 + 4x^3 + \dots,$$

очевидно, сходится при $x = 0$ и расходится при $x = 1$.

Разумная постановка вопроса должна быть такова: при каких x сходится ряд (1) и при каких он расходится? Оказывается, что множество тех x, при которых сходится степенной ряд (1), всегда имеет очень простое строение: это — промежуток, симметричный относительно точки $x = 0$. Докажем этот замечательный факт (открытый Н. Г. Абелем, 1802 — 1829).

Лемма Абеля. *Если степенной ряд* (1) *сходится при некотором значении* $x_0 \neq 0$, *то он абсолютно сходится при каждом* x, *у которого*

$$|x| < |x_0|. \tag{6}$$

В самом деле, так как общий член сходящегося ряда

$$c_0 + c_1 x_0 + c_2 x_0^2 + c_3 x_0^3 + \dots$$

стремится к нулю с возрастанием своего номера, то

$$\lim_{n \to \infty} c_n x_0^n = 0.$$

Но если какая-нибудь последовательность имеет предел, то множество членов этой последовательности ограничено. Поэтому существует такое число M, что

$$|c_n x_0^n| < M \quad (n = 0, 1, 2, \ldots). \tag{7}$$

Заметив это, рассмотрим ряд (1) при значении x, удовлетворяющем неравенству (6). Этот ряд можно переписать так:

$$c_0 + c_1 x_0 \frac{x}{x_0} + c_2 x_0^2 \left(\frac{x}{x_0}\right)^2 + c_3 x_0^3 \left(\frac{x}{x_0}\right)^3 + \ldots \tag{8}$$

В силу (7) ряд, состоящий из абсолютных величин членов последнего ряда, имеет мажорантный ряд

$$M + M\left|\frac{x}{x_0}\right| + M\left|\frac{x}{x_0}\right|^2 + M\left|\frac{x}{x_0}\right|^3 + \ldots,$$

а так как этот ряд сходится (ибо это есть геометрическая прогрессия со знаменателем $\left|\frac{x}{x_0}\right| < 1$), то ряд (8), или, что то же самое, ряд (1), сходится абсолютно. Лемма доказана.

Рассмотрим теперь какой-нибудь степенной ряд (1) и образуем множество A, состоящее из тех x, при которых наш ряд сходится. Это множество заведомо не пусто, ибо всякий ряд (1) сходится при $x = 0$.

Пусть, далее, B есть множество абсолютных величин всех x, входящих в A, и R — точная верхняя граница[1]) этого множества B

$$R = \sup B.$$

Различим три случая, которые могут иметь место:

$$1)\ R = 0, \quad 2)\ 0 < R < +\infty, \quad 3)\ R = +\infty.$$

В первом из этих случаев ряд (1) сходится т о л ь к о при $x = 0$. Действительно, если он сходится при некотором x, то $x \in A$, откуда $|x| \in B$ и $|x| \leqslant R = 0$.

Рассмотрим теперь случай 2). В этом случае ряд (1) сходится, и притом абсолютно, при каждом x, входящем в о т к р ы т ы й промежуток $(-R, +R)$. В точках же, лежащих вне з а м к н у т о г о промежутка $[-R, +R]$, ряд расходится. В самом деле, пусть сначала $x \in (-R, +R)$. Тогда $|x| < R$ и по свойствам точной верхней границы в множестве B обязательно найдётся число, большее чем $|x|$. Это число (по самому определению множества B) является абсолютной величиной какого-то элемента x_0 множества A. Иначе говоря, $|x| < |x_0|$, причём $x_0 \in A$. Но тогда по лемме Абеля ряд (1) абсолютно сходится для нашего x.

---

[1]) Если B — множество, не ограниченное сверху, то, как обычно, полагаем $R = +\infty$.

Если же x лежит вне $[-R, +R]$, то $|x| > R$. Значит, $|x|$ не входит в B, а сам x не входит в A, т. е. при этом x ряд расходится.

Наконец, в случае 3) множество B не ограничено сверху. Значит, для любого действительного x найдётся элемент множества B, больший чем $|x|$. Этот элемент, как и выше, можно записать в виде $|x_0|$, где $x_0 \in A$. В силу леммы Абеля можно утверждать, что ряд (1) абсолютно сходится при взятом значении x. А так как это было произвольное действительное число, то ряд оказывается абсолютно сходящимся на всей оси.

Подводя итог нашему исследованию, мы можем формулировать его в виде следующей теоремы.

Теорема. *Всякому ряду* (1) *отвечает такое неотрицательное число*[1]) R, *что ряд абсолютно сходится в открытом промежутке* $(-R, +R)$ *и расходится вне замкнутого промежутка* $[-R, +R]$.

К этой теореме надлежит сделать ряд добавочных замечаний:

1) Смысл теоремы состоит в том, что множество тех значений x, при которых сходится степенной ряд (1), всегда есть некоторый промежуток, симметричный относительно точки $x=0$. Этот промежуток называется *промежутком сходимости* ряда.

2) Промежуток сходимости может вырождаться в точку $x=0$, что имеет место при $R=0$. Возможны также ряды (когда $R=+\infty$), у которых промежутком сходимости служит вся числовая ось.

3) В теореме совершенно ничего не говорится о том, как ведёт себя ряд на концах промежутка сходимости, т. е. в точках $x=-R$, $x=+R$. Как мы увидим, различные ряды в этих точках ведут себя по-разному, и потому никаких общих утверждений по этому поводу высказать нельзя.

4) Мы уже говорили, что ряд вида (2), т. е. расположенный по степеням $x-a$, приводится к виду (1) подстановкой $x-a=x'$. Поэтому всё сказанное выше переносится и на ряды (2). Промежуток сходимости такого ряда будет симметричен относительно точки $x=a$ и будет иметь концами точки $a-R$ и $a+R$.

Приведём теперь ряд примеров, которые покажут, что все рассмотренные нами случаи могут действительно реализоваться на практике.

Пример 1. Ряд

$$\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\dots$$

сходится на всей числовой оси. В самом деле, применяя к нему признак Даламбера (теорема 3 из п°37), будем иметь (при любом $x \neq 0$)

$$a_n=\frac{x^n}{n!}, \quad a_{n+1}=\frac{x^{n+1}}{(n+1)!}, \quad \frac{a_{n+1}}{a_n}=\frac{x}{n+1} \xrightarrow[n\to\infty]{} 0,$$

и ряд сходится.

[1]) Это число называется *радиусом сходимости* ряда.

Пример 2. Ряд

$$1!\,x+2!\,x^2+3!\,x^3+\ldots$$

сходится только при $x=0$. Действительно, если $x\neq 0$, то

$$a_n=n!\,x^n,\quad a_{n+1}=(n+1)!\,x^{n+1},\quad \frac{a_{n+1}}{a_n}=(n+1)\,x\longrightarrow\infty,$$

и ряд расходится.

Пример 3. Прогрессия

$$1+x+x^2+x^3+\ldots$$

сходится при $|x|<1$ и расходится при $|x|\geqslant 1$. Значит, промежутком её сходимости является открытый промежуток $(-1,+1)$. На его концах прогрессия расходится.

Пример 4. Применяя признак Даламбера к ряду

$$x+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\frac{x^4}{4^2}+\ldots, \tag{9}$$

получим

$$a_n=\frac{x^n}{n^2},\quad a_{n+1}=\frac{x^{n+1}}{(n+1)^2},\quad \frac{a_{n+1}}{a_n}=\left(\frac{n}{n+1}\right)^2 x\longrightarrow x.$$

Значит, ряд абсолютно сходится при $|x|<1$ и расходится при $|x|>1$. Если же $x=1$, то наш ряд принимает вид

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\ldots \tag{10}$$

В n° 35 было установлено, что этот ряд сходится. При $x=-1$ получается ряд, для которого (10) служит рядом абсолютных величин. Поэтому промежутком сходимости ряда (9) служит замкнутый промежуток $[-1,+1]$. На концах этого промежутка ряд (9) сходится абсолютно.

Пример 5. Аналогичное исследование, применённое к ряду

$$x^2-\frac{x^4}{2}+\frac{x^6}{3}-\frac{x^8}{4}+\ldots,$$

показывает, что он имеет тот же промежуток сходимости $[-1,+1]$, что и ряд (9), но на обоих концах этого промежутка сходимость будет неабсолютной.

Пример 6. Ряд

$$x+\frac{x^2}{2}+\frac{x^3}{3}+\frac{x^4}{4}+\ldots$$

сходится при $-1\leqslant x<+1$. Значит, его промежуток сходимости $[-1,+1)$ содержит свой левый конец $x=-1$ (причём в этом конце сходимость неабсолютная) и не содержит конца правого, где ряд расходится.

Во всех рассмотренных примерах, в которых было $0 < R < +\infty$, оказывалось $R = +1$. Достаточно, однако, в любом из них взять вместо x отношение $\frac{x}{a}$ $(a > 0)$, чтобы получить пример ряда, у которого $R = a$.

40. Свойства суммы степенного ряда. Рассмотрим степенной ряд

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \ldots, \tag{11}$$

радиус сходимости R которого отличен от нуля (случай $R = +\infty$ мы не исключаем). Сумма этого ряда будет функцией аргумента x, определённой во всех точках промежутка сходимости. Обозначим эту функцию через $f(x)$ и займёмся изучением её свойств.

Теорема 1. *Функция $f(x)$ непрерывна во всех точках, лежащих внутри промежутка сходимости ряда.*

Подчеркнём, что, говоря о точках, лежащих внутри промежутка сходимости, мы исключаем его концы $\pm R$, хотя ряд может сходиться и в этих точках [1]).

Переходя к доказательству, выберем какую-либо точку x_0, лежащую внутри промежутка сходимости $-R < x_0 < R$, и будем устанавливать непрерывность $f(x)$ именно в этой точке. Для этой цели выберем и закрепим ещё одну точку X, для которой выполняются неравенства

$$-X < x_0 < X, \quad 0 < X < R.$$

Эта точка тоже содержится в открытом интервале $(-R, +R)$, и потому наш ряд (11) абсолютно сходится в ней. Иными словами, сходится ряд

$$|c_0| + |c_1| X + |c_2| X^2 + |c_3| X^3 + \ldots \tag{12}$$

Теперь возьмём совершенно произвольную точку x из промежутка $[-X, X]$ (в частности, это может быть и точка x_0) и составим ряд

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \ldots \tag{13}$$

Пусть $\overline{R}_n$ и $R_n(x)$ суть суммы остатков (после n-го члена) рядов (12) и (13), т. е.

$$\overline{R}_n = |c_{n+1}| X^{n+1} + |c_{n+2}| X^{n+2} + \ldots,$$
$$R_n(x) = c_{n+1} x^{n+1} + c_{n+2} x^{n+2} + \ldots$$

Нетрудно видеть [1]), что

$$|R_n(x)| \leqslant \overline{R}_n. \tag{14}$$

---

[1]) Можно доказать, что в случае сходимости ряда в одной из точек $\pm R$ функция $f(x)$ остаётся непрерывной в этой точке, но мы не будем останавливаться на этом.

Заметив это, возьмём произвольное $\varepsilon > 0$ и найдём столь большое n, чтобы оказалось

$$\overline{R}_n < \frac{\varepsilon}{3}.$$

Такое n существует, ибо сумма остатка сходящегося ряда (12) стремится к нулю.

В силу (14) при этом n и при всех x из $[-X, X]$ окажется

$$|R_n(x)| < \frac{\varepsilon}{3}.$$

В частности, и $|R_n(x_0)| < \frac{\varepsilon}{3}$. Положим, далее,

$$S_n(x) = c_0 + c_1 x + \ldots + c_n x^n.$$

Тогда

$$f(x) = S_n(x) + R_n(x), \quad f(x_0) = S_n(x_0) + R_n(x_0).$$

Значит,

$$|f(x) - f(x_0)| \leqslant |S_n(x) - S_n(x_0)| + |R_n(x)| + |R_n(x_0)|,$$

и тем более

$$|f(x) - f(x_0)| < |S_n(x) - S_n(x_0)| + \frac{2}{3}\varepsilon. \qquad (15)$$

Ещё раз подчеркнём, что это неравенство доказано для всех x из $[-X, X]$.

Отметим теперь, что $S_n(x)$ есть обыкновенный алгебраический многочлен. Значит, это — функция непрерывная. Поэтому для нашего ε существует такое $\delta > 0$, что, как только $|x - x_0| < \delta$, так сейчас же

$$|S_n(x) - S_n(x_0)| < \frac{\varepsilon}{3}. \qquad (16)$$

Если же $|x - x_0| < \delta$ и $x \in [-X, X]$, то из (15) и (16) вытекает, что

$$|f(x) - f(x_0)| < \varepsilon.$$

Этим и доказана непрерывность $f(x)$ в точке x_0.

Теорема 2. *Равенство*

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \ldots \qquad (17)$$

---

[1]) Чтобы доказать это с полной строгостью, достаточно в очевидном неравенстве

$$\left|\sum_{k=n+1}^{n+m} c_k x^k\right| \leqslant \sum_{k=n+1}^{n+m} |c_k| X^k$$

перейти к пределу при $m \longrightarrow +\infty$.

можно почленно интегрировать по любому замкнутому промежутку $[a, b]$, *содержащемуся в открытом промежутке* $(-R, +R)$, *где* R — *радиус сходимости ряда* (11).

Иными словами, мы утверждаем, что для всякого промежутка $[a, b]$, у которого $-R < a < b < R$ будет

$$\int_a^b f(x)\,dx = c_0(b-a) + c_1\frac{b^2-a^2}{2} + c_2\frac{b^3-a^3}{3} + \dots \qquad (18)$$

Доказательство этой теоремы сходно с доказательством теоремы 1. Именно, как и выше, мы начинаем с того, что закрепляем точку X, удовлетворяющую неравенствам

$$-R < -X \leqslant a < b \leqslant X < R.$$

Сохраняя введённые выше обозначения, мы попрежнему будем иметь, что при любом натуральном n и любом $x \in [-X, X]$ будет

$$|R_n(x)| \leqslant \overline{R}_n.$$

Поэтому при всех x из $[-X, X]$ и тем более при всех x из $[a, b]$ будет

$$|f(x) - S_n(x)| \leqslant \overline{R}_n.$$

В таком случае (по теореме 11 из п° 23)

$$\left|\int_a^b \{f(x) - S_n(x)\}\,dx\right| \leqslant \overline{R}_n(b-a),$$

или, что то же самое,

$$\left|\int_a^b f(x)\,dx - \int_a^b S_n(x)\,dx\right| \leqslant \overline{R}_n(b-a).$$

Правая часть этого неравенства с возрастанием n стремится к нулю. Значит,

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\int_a^b S_n(x)\,dx.$$

Остаётся заметить, что

$$\int_a^b S_n(x)\,dx = c_0(b-a) + c_1\frac{b^2-a^2}{2} + \dots + c_n\frac{b^{n+1}-a^{n+1}}{n+1},$$

так что этот интеграл есть не что иное, как частичная сумма ряда, стоящего в (18) справа.

Теорема 3. *Равенство* (17) *можно почленно дифференцировать внутри промежутка сходимости ряда* (11).

Иначе говоря, сумма $f(x)$ ряда (11) в каждой точке x из открытого промежутка $(-R, +R)$ имеет производную $f'(x)$, и справедливо равенство

$$f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + 4c_4 x^3 + \ldots$$

Доказательство этой теоремы сравнительно сложно, и мы предпошлём ему две леммы.

Лемма 1. *Если* $0 < q < 1$, *то ряд*

$$q + 2q^2 + 3q^3 + 4q^4 + \ldots \tag{19}$$

сходится.

Эта лемма устанавливается с помощью признака Даламбера: здесь

$$a_n = nq^n, \quad a_{n+1} = (n+1)\,q^{n+1},$$

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{n}\,q \xrightarrow[n\to\infty]{} q < 1.$$

Лемма 2. *Степенные ряды*

$$c_1 x + c_2 x^2 + c_3 x^3 + \ldots, \tag{20}$$

$$c_1 x + 2c_2 x^2 + 3c_3 x^3 + \ldots \tag{21}$$

имеют равные радиусы сходимости.

Обозначим упомянутые радиусы сходимости рядов (20) и (21) соответственно через R и R'.

Предположим, что $R' > 0$ и пусть $0 < z < R'$. Тогда в точке z ряд (21) сходится абсолютно. Иначе говоря, сходится положительный ряд

$$|c_1|\,z + 2\,|c_2|\,z^2 + 3\,|c_3|\,z^3 + \ldots$$

Тем более сходится ряд

$$|c_1|\,z + |c_2|\,z^2 + |c_3|\,z^3 + \ldots$$

Это показывает, что в точке z сходится (даже абсолютно) ряд (20) и потому

$$z \leqslant R. \tag{22}$$

Здесь z подчинено единственному условию $0 < z < R'$. Значит, можно устремить в (22) z к R' и перейти к пределу.

Таким образом,

$$R' \leqslant R. \tag{23}$$

Это неравенство, установленное нами для $R' > 0$, очевидным образом справедливо и при $R' = 0$.

Установим теперь, что

$$R \leqslant R'. \tag{24}$$

При этом можно предполагать, что $R > 0$, ибо иначе неравенство (24) очевидно. Пусть $0 < z < R$. Покажем, что ряд (21) абсолютно сходится в этой точке. С этой целью возьмём ещё одну точку y, удовлетворяющую неравенствам $z < y < R$. Ряд (20) в этой точке сходится и потому его общий член

стремится к нулю. Тем более этот общий член ограничен. Значит, существует такое постоянное число M, что при всех натуральных n будет

$$|c_n y^n| < M. \tag{25}$$

Заметив это, перепишем ряд

$$|c_1|z + 2|c_2|z^2 + 3|c_3|z^3 + \dots \tag{26}$$

так:

$$|c_1 y|\frac{z}{y} + 2|c_2 y^2|\left(\frac{z}{y}\right)^2 + 3|c_3 y^3|\left(\frac{z}{y}\right)^3 + \dots$$

В силу (25) члены этого последнего ряда меньше, чем соответствующие члены ряда

$$Mq + 2Mq^2 + 3Mq^3 + \dots, \tag{27}$$

где положено для краткости $q = \frac{z}{y}$. По лемме 1 ряд (27) сходится, а потому сходится и ряд (26).

Итак, действительно, ряд (21) сходится в точке z. Но тогда

$$z \leqslant R',$$

откуда, как и выше, переходя к пределу при z, стремящемся к R, мы и получаем неравенство (24). Лемма доказана.

Из этой леммы вытекает важное

Следствие. *Степенной ряд* (11) *и ряд*

$$c_1 + 2c_2x + 3c_3x^2 + \dots, \tag{28}$$

полученный из него почленным дифференцированием, имеют одинаковые радиусы[1] *сходимости.*

В самом деле, ряд (11) имеет тот же радиус сходимости, что и его остаток (20). Тот же радиус по лемме имеет и ряд (21), который получается из (28) умножением всех членов на x. Остаётся показать, что (21) и (28) имеют равные радиусы сходимости. Но у них даже и промежутки сходимости одинаковы. Действительно, если ряд (28) сходится при каком-либо x, то, умножая его члены на x, мы не нарушим сходимости; если же ряд (28) при каком-либо x расходится, то очевидно $x \neq 0$, а тогда, умножая члены ряда (28) на это x, мы не сможем придти к сходящемуся ряду.

Теперь мы можем, наконец, вернуться к теореме 3. Пусть имеет место равенство (17), причём радиус сходимости R ряда, стоящего в правой части (17), положителен, $R > 0$. Дифференцируя этот ряд почленно, мы приходим к ряду (28), имеющему тот же радиус сходимости R. Обозначим сумму этого ряда через $\varphi(x)$. По теореме 1 это есть функция, непрерывная в $(-R, +R)$. Закрепим какую-либо точку z этого интервала и рассмотрим замкнутый промежуток[2] $[0, z]$.

---

[1]) Не следует думать, что и промежутки сходимости этих рядов тождественны. Например, ряд $\sum \frac{x^n}{n^2}$ имеет замкнутый промежуток сходимости $[-1, +1]$, в то время как продифференцированный ряд в точке $x = 1$ расходится.

[2]) Мы считаем $z > 0$. Если $z < 0$, то надо ввести промежуток $[z, 0]$, а в остальном все рассуждения сохраняются, ибо при перестановке пределов интегрирования интеграл меняет знак.

Этот промежуток содержится в $(-R, +R)$ и (на основании теоремы 2) можно почлеино проинтегрировать равенство

$$\varphi(x) = c_1 + 2c_2x + 3c_3x^2 + \dots$$

по промежутку $[0, z]$. Таким образом,

$$\int_0^z \varphi(x)\,dx = c_1z + c_2z^2 + c_3z^3 + \dots$$

Сравнивая это равенство с (17), находим

$$f(z) = c_0 + \int_0^z \varphi(x)\,dx.$$

По теореме п° 24 о дифференцировании определённого интеграла по верхнему пределу правая часть последнего равенства имеет производную, равную $\varphi(z)$. Но тогда ту же производную имеет и его левая часть. Стало быть, $f'(z)$ существует и $f'(z) = \varphi(z)$. Заменяя букву z на x и вспоминая, что $\varphi(x)$ есть сумма ряда (28), мы завершаем доказательство.

Так как мы можем применить доказанную теорему и к продифференцированному ряду, а затем снова применить её же и т. д., то справедлива

Теорема 4. *Степенной ряд можно дифференцировать почленно внутри его промежутка сходимости любое число раз.*

Эту же теорему можно формулировать и так: если ряд, стоящий в равенстве

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \dots,$$

имеет радиус сходимости $R > 0$, то $f(x)$ имеет производные всех порядков, причём для всех $x \in (-R, +R)$ будет

$$\begin{aligned} f'(x) &= c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 + \dots, \\ f''(x) &= 2c_2 + 3 \cdot 2c_3x + 4 \cdot 3c_4x^2 + \dots, \\ f'''(x) &= 3 \cdot 2 \cdot 1c_3 + 4 \cdot 3 \cdot 2c_4x + \dots, \end{aligned}$$

и все написанные здесь ряды имеют тот же радиус сходимости R.

41. Разложение логарифма и составление таблиц логарифмов. Применим изложенные результаты к одному вопросу, имеющему фундаментальное значение для курса элементарной алгебры. Мы имеем в виду вопрос о том, как составляются логарифмические таблицы.

Рассмотрим геометрическую прогрессию

$$1 - x + x^3 - x^3 + \dots$$

Она сходится, когда её знаменатель $q = -x$ по абсолютной величине меньше единицы, т. е. когда $-1 < x < 1$, причём сумма её в этом открытом промежутке равна $\frac{1}{1+x}$.

Итак,

$$\frac{1}{1+x}=1-x+x^2-x^3+\ldots$$

Проинтегрируем это равенство почленно по промежутку $[0, z]$, где $-1<z<1$, что законно по теореме 2 из п° 40. В результате мы приходим к равенству

$$\ln(1+z)=z-\frac{z^2}{2}+\frac{z^3}{3}-\frac{z^4}{4}+\ldots,$$

которое можно записать и так:

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots \qquad (29)$$

Это равенство доказано нами для $-1<x<1$. Для $x=-1$ нельзя и ставить вопроса о справедливости этого равенства, ибо обе его части теряют при этом x числовой смысл (слева $\ln 0$, справа расходящийся ряд). Напротив, при $x=+1$ и слева и справа получаются выражения, имеющие числовой смысл, но мы пока всё же ещё не имеем оснований утверждать справедливость равенства этих выражений, т. е. что

$$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots \qquad (30)$$

Докажем, что равенство (30) всё же верно. С этой целью будем исходить не из бесконечного ряда

$$1-x+x^2-x^3+\ldots,$$

а из конечной суммы

$$1-x+x^2-x^3+x^4-x^5+\ldots+x^{2n-2}-x^{2n-1}=\frac{1-x^{2n}}{1+x} \qquad (x\neq -1).$$

Отсюда

$$\frac{1}{1+x}=(1-x+x^2-x^3+\ldots+x^{2n-2}-x^{2n-1})+\frac{x^{2n}}{1+x},$$

или, интегрируя от 0 до 1,

$$\ln 2=\int_0^1\frac{dx}{1+x}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{1}{2n-1}-\frac{1}{2n}+\int_0^1\frac{x^{2n}}{1+x}\,dx.$$

Так как $\frac{x^{2n}}{1+x}<x^{2n}$ при $x>0$, то

$$\int_0^1\frac{x^{2n}}{1+x}\,dx<\int_0^1 x^{2n}\,dx=\frac{1}{2n+1};$$

следовательно,

$$\lim_{n\to\infty}\int_0^1 \frac{x^{2n}}{1+x}\,dx=0 \quad \text{и} \quad \ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

Равенство (30) доказано [1]). Значит, равенство (29) верно для $-1<x\leqslant 1$.

Формула (29) является исходной при составлении таблицы логарифмов. Однако непосредственное её использование для этой цели затрудняется [2]) тем обстоятельством, что она справедлива лишь при $-1<x\leqslant 1$. Поэтому приходится несколько преобразовать эту формулу.

Заменяя в (29) x на $-x$, получим для $-1\leqslant x<1$

$$\ln(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-\frac{x^4}{4}-\ldots \qquad (31)$$

Вычитая это равенство из (29), находим:

$$\ln\frac{1+x}{1-x}=2\left(x+\frac{x^3}{3}+\frac{x^5}{5}+\ldots\right).$$

Положим здесь

$$x=\frac{1}{2N+1},$$

где N — натуральное число. Замечая, что при этом $\frac{1+x}{1-x}=\frac{N+1}{N}$,

---

[1]) Дадим ещё одно красивое доказательство равенства (30). Положим $S_{2n}=1-\frac{1}{2}+\frac{1}{3}-\ldots-\frac{1}{2n}$. Прибавим к правой части и вычтем из неё сумму $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}$, причём, прибавляя, запишем эту сумму в виде $2\left(\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2n}\right)$. Очевидно, в результате этих операций мы получим:

$$S_{2n}=\left(1+\frac{1}{2}+\ldots+\frac{1}{2n}\right)-\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right)=\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{2n}.$$

Отсюда $S_{2n}=\sum_{k=1}^{n}\frac{1}{1+\frac{k}{n}}\cdot\frac{1}{n}$. Это — интегральная сумма, стремящаяся при $n\to\infty$ к $\int_0^1\frac{dx}{1+x}=\ln 2$, чем снова доказано (30).

[2]) При помощи некоторых искусственных приёмов эту трудность можно было бы обойти. Например, мы могли бы вычислять $\ln 2$ не при помощи очень медленно сходящегося ряда (30), а положив в (29) $x=-\frac{1}{2}$, что дало бы $-\ln 2=-\frac{1}{2}-\frac{1}{2\cdot 2^2}-\frac{1}{3\cdot 2^3}-\frac{1}{4\cdot 2^4}-\ldots$

Здесь для получения какой-либо заранее заданной точности пришлось бы брать гораздо меньше членов, чем в ряде (30).

Положив, далее, в (29) $x=-\frac{1}{3}$, мы нашли бы $\ln\frac{2}{3}$, что привело бы нас и к $\ln 3$ (поскольку $\ln 2$ уже известен), и т. д.

мы получаем основную для интересующего нас вопроса формулу

$$\ln(N+1)=\ln N+\frac{2}{2N+1}+\frac{2}{3(2N+1)^3}+\frac{2}{5(2N+1)^5}+\cdots \quad (32)$$

Эта формула позволяет нам находить $\ln(N+1)$, если уже известен $\ln N$. Но так как $\ln 1=0$, то, полагая в (32) последовательно $N=1$, $N=2$, $N=3$, ... , мы сможем получить из этой формулы один за другим логарифмы всех натуральных чисел до любого интересующего нас предела.

Чтобы производимые при этом вычисления были надёжными, надо уметь оценивать ошибку [1]), получающуюся при замене суммы ряда, стоящего в (32), на его частичную сумму. Для этого, очевидно, надо оценить сумму остатка этого ряда, т. е.

$$\rho_p(N)=\frac{2}{(2p+1)(2N+1)^{2p+1}}+\frac{2}{(2p+3)(2N+1)^{2p+3}}+$$
$$+\frac{2}{(2p+5)(2N+1)^{2p+5}}+\cdots$$

Ясно, что

$$0<\rho_p(N)<\frac{2}{(2p+1)(2N+1)^{2p+1}}+\frac{2}{(2p+1)(2N+1)^{2p+3}}+$$
$$+\frac{2}{(2p+1)(2N+1)^{2p+5}}+\cdots$$

Суммируя написанную здесь геометрическую прогрессию, находим

$$0<\rho_p(N)<\frac{1}{2N(N+1)}\cdot\frac{1}{2p+1}\cdot\frac{1}{(2N+1)^{2p-1}}. \quad (33)$$

Подчеркнём, что $\rho_p(N)$ есть ошибка приближённого равенства

$$\ln(N+1)=\ln N+\frac{2}{2N+1}+\frac{2}{3(2N+1)^3}+\cdots+$$
$$+\frac{2}{(2p-1)(2N+1)^{2p-1}}. \quad (34)$$

Применим, например, формулу (34) для вычисления $\ln 2$. Для этого надо положить $N=1$. Оценка (33) при этом принимает вид

$$0<\rho_p(1)<\frac{1}{4(2p+1)3^{2p-1}}.$$

---

[1]) Ошибкой приближённого равенства мы раз навсегда условимся называть то число, которое следует прибавить к приближённому значению величины, чтобы получить точное. Иными словами, это есть поправка к приближённому значению. Например, ошибка равенства $\frac{1}{3}=0{,}333$ есть $0{,}000(3)$, а ошибка равенства $\frac{2}{3}=0{,}667$ есть $-0{,}000(3)$.

Возьмём $p=6$. Так как

$$\frac{1}{52\cdot 3^{11}}<0{,}00000012,$$

то ошибка $\rho_6(1)$ равенства

$$\ln 2=\frac{2}{3}+\frac{2}{3\cdot 3^3}+\frac{2}{5\cdot 3^5}+\frac{2}{7\cdot 3^7}+\frac{2}{9\cdot 3^9}+\frac{2}{11\cdot 3^{11}} \qquad (35)$$

удовлетворяет неравенству

$$0<\rho_6(1)<0{,}00000012.$$

Запишем каждый член правой части (35) в форме десятичной дроби с 8-ю верными знаками. Это означает, что абсолютная величина поправки, которую надо прибавить к десятичному разложению числа, чтобы получить точное значение этого числа, не больше, чем

$$0{,}000000005.$$

Чтобы сделать наши оценки более точными, мы будем, оканчивая запись десятичной дроби, ставить в скобках знак упомянутой поправки [1]). Таким образом,

$$\frac{2}{3}=0{,}66666667\ (-),$$

$$\frac{2}{3\cdot 3^3}=0{,}02469136\ (-),$$

$$\frac{2}{5\cdot 3^5}=0{,}00164609\ (+),$$

$$\frac{2}{7\cdot 3^7}=0{,}00013064\ (+),$$

$$\frac{2}{9\cdot 3^9}=0{,}00001129\ (+),$$

$$\frac{2}{11\cdot 3^{11}}=0{,}00000103\ (-).$$

Сумма выписанных здесь десятичных дробей равна

$$0{,}69314708. \qquad (36)$$

Это число не является, однако, точным значением суммы, стоящей в (35) справа. Если от каждой дроби, отмеченной знаком (—), мы отнимем 0,000000005, а дроби, отмеченные знаком (+), оставим без изменения, то полученная от сложения этих дробей сумма будет меньше суммы, стоящей в (35) справа. Но знаком (—) у нас отмечены 3 дроби. Поэтому, отнимая от (36) 0,000000015, мы получим число, меньшее суммы, стоящей в правой части (35). При-

[1]) То-есть у дробей избыточных писать (—), а у недостаточных (+).

меняя аналогичные соображения к дробям, отмеченным знаком (+), а также учитывая оценку для $\rho_6(1)$, получим окончательно:

$$0{,}693147065 < \ln 2 < 0{,}693147215, \tag{37}$$

откуда

$$\ln 2 = 0{,}693147\ (+) \tag{38}$$

и все 6 знаков здесь верны.

По всей вероятности, основные принципы, при помощи которых вычисляются логарифмы, достаточно выяснены на рассмотренном примере. Однако, ввиду чрезвычайно большого значения этого вопроса именно для элементарной математики, мы фактически составим таблицу *десятичных* [1]) логарифмов для чисел от 1 до 10.

Подставляя в (32) $N=2$, находим:

$$\ln 3 = \ln 2 + \frac{2}{5} + \frac{2}{3\cdot 5^3} + \frac{2}{5\cdot 5^5} + \dots \tag{39}$$

Оценка (33) здесь принимает вид

$$0 < \rho_p(2) < \frac{1}{12}\cdot\frac{1}{(2p+1)\cdot 5^{2p-1}}.$$

В частности,

$$0 < \rho_4(2) < \frac{1}{108\cdot 5^7} < 0{,}00000013. \tag{40}$$

С другой стороны,

$$\frac{2}{5} = 0{,}40000000,$$

$$\frac{2}{3\cdot 5^3} = 0{,}00533333\ (+),$$

$$\frac{2}{5\cdot 5^5} = 0{,}00012800,$$

$$\frac{2}{7\cdot 5^7} = 0{,}00000366\ (-).$$

Отсюда

$$\frac{2}{5} + \frac{2}{3\cdot 5^3} + \frac{2}{5\cdot 5^5} + \frac{2}{7\cdot 5^7} = 0{,}40546499,$$

причём абсолютная величина ошибки этого равенства меньше, чем 0,000000005. Сопоставляя это с (39), (40) и (37), находим:

$$1{,}098612050 < \ln 3 < 1{,}098612340. \tag{41}$$

Отсюда получается значение $\ln 3$ с 6-ю верными знаками

$$\ln 3 = 1{,}098612\ (+). \tag{42}$$

Что касается $\ln 4$, то его можно было бы найти тем же способом, но проще получить его из (37) на основании равенства $\ln 4 = 2\ln 2$. Это сразу даёт, что

$$1{,}386294130 < \ln 4 < 1{,}386294430, \tag{43}$$

откуда с 6-ю верными знаками

$$\ln 4 = 1{,}386294\ (+).$$

[1]) До сих пор речь шла о логарифмах натуральных.

Подставляя в формулу (32) $N=4$, получим:

$$\ln 5=\ln 4+\frac{2}{9}+\frac{2}{3\cdot 9^3}+\frac{2}{5\cdot 9^5}+\dots$$

Оценка (33) даёт:

$$0<\rho_3(4)<\frac{1}{280\cdot 9^5}<0{,}0000001. \tag{44}$$

Далее,

$$\frac{2}{9}=0{,}22222222\ (+),$$

$$\frac{2}{3\cdot 9^3}=0{,}00091449\ (+),$$

$$\frac{2}{5\cdot 9^5}=0{,}00000677\ (+),$$

откуда

$$\frac{2}{9}+\frac{2}{3\cdot 9^3}+\frac{2}{5\cdot 9^5}=0{,}22314348,$$

причём ошибка здесь положительна, но меньше, чем 0,000000015. Сопоставляя это с (43) и (44), находим:

$$1{,}609437610<\ln 5<1{,}609438025. \tag{45}$$

Остановимся теперь на вопросе о переходе от натуральных логарифмов к десятичным.

Если N — произвольное положительное число, а $\lg N$ — его десятичный логарифм, то

$$N=10^{\lg N}.$$

Логарифмируя это равенство по основанию e, получим:

$$\ln N=\lg N\cdot\ln 10.$$

Стало быть,

$$\lg N=\frac{\ln N}{\ln 10}.$$

Вычислим $\ln 10$, опираясь на соотношения (37) и (45) и на то, что $\ln 10=\ln 2+\ln 5$. Очевидно,

$$2{,}302584675<\ln 10<2{,}302585240$$

и тем более

$$2{,}302584<\ln 10<2{,}302586. \tag{46}$$

Теперь мы уже в состоянии вычислить $\lg 2$ и $\lg 3$. Для того чтобы получаемые результаты были абсолютно надёжными, мы установим для этих логарифмов сначала границы, между которыми они лежат [1]).

---

[1]) Мы преследуем цель получить интересующие нас логарифмы не с очень большой, но совершенно гарантированной точностью. Обычно вычисления ведут более экономными способами, которые, однако, иногда оставляют сомнительным последний знак, даже если вычисляется целый ряд лишних десятичных знаков. Тем не менее, мы вовсе не имеем в виду рекомендовать замену обычной методики вычислений на применяемую нами.

Именно, из неравенств (37) и (46) следует, что

$$\frac{0{,}693147}{2{,}302586} < \lg 2 < \frac{0{,}693148}{2{,}302584}.$$

Обратим крайние члены этого двойного неравенства в десятичные дроби, причём для нижней границы возьмём эту дробь с недостатком, а для верхней — с избытком. В результате получаются неравенства

$$0{,}301029 < \lg 2 < 0{,}3010305. \tag{47}$$

Умножая это двойное неравенство на 2 и на 3, находим:

$$0{,}602058 < \lg 4 < 0{,}602061, \tag{48}$$

$$0{,}903087 < \lg 8 < 0{,}9030915. \tag{49}$$

Аналогично из (41) и (46) находим сначала, что

$$\frac{1{,}098612}{2{,}302586} < \lg 3 < \frac{1{,}098613}{2{,}302584},$$

а затем

$$0{,}477120 < \lg 3 < 0{,}477122. \tag{50}$$

Удваивая, находим:

$$0{,}954240 < \lg 9 < 0{,}954244. \tag{51}$$

Складывая неравенства (47) и (50), получаем:

$$0{,}778149 < \lg 6 < 0{,}7781525. \tag{52}$$

Вычитая же все члены неравенства (47) из единицы, находим:

$$0{,}6989695 < \lg 5 < 0{,}698971. \tag{53}$$

У нас уже найдены десятичные логарифмы всех чисел первого десятка, кроме $\lg 7$.

Для нахождения этого последнего логарифма мы снова должны обратиться к формуле (32), положив в ней $N = 6$. Это даёт

$$\ln 7 = \ln 6 + \frac{2}{13} + \frac{2}{3 \cdot 13^3} + \frac{2}{5 \cdot 13^5} + \dots$$

Из (37) и (41) мы находим, что

$$1{,}7917591 < \ln 6 < 1{,}7917596.$$

Кроме того, согласно (38) имеем:

$$0 < \rho_3(6) < \frac{1}{84} \cdot \frac{1}{5 \cdot 13^3} < 0{,}0000011.$$

Наконец,

$$\frac{2}{13} = 0{,}1538462\ (-),$$

$$\frac{2}{3 \cdot 13^3} = 0{,}0003034\ (+),$$

откуда с ошибкой, абсолютно меньшей 0,00000005,

$$\frac{2}{13} + \frac{2}{3 \cdot 13^3} = 0{,}1541496.$$

Таким образом,

$$1{,}94590865 < \ln 7 < 1{,}94591035.$$

Отсюда и из (46) получаем сначала, что

$$\frac{1{,}945908}{2{,}302586} < \lg 7 < \frac{1{,}945911}{2{,}302584},$$

а затем, что

$$0{,}845096 < \lg 7 < 0{,}845099. \tag{54}$$

Собирая формулы (47) — (54) и присоединяя к ним тривиальные тождества $\lg 1 = 0$, $\lg 10 = 1$, получаем таблицу:

N	$\lg N$	N	$\lg N$
1	0,00000	6	0,77815
2	0,30103	7	0,84510
3	0,47712	8	0,90309
4	0,60206	9	0,95424
5	0,69897	10	1,00000

Все написанные здесь знаки верны!

В заключение мы ещё раз хотим предостеречь (см. подстрочное примечание на стр. 463) читателя от мысли, что таблицы логарифмов в действительности вычислялись в точности так, как это было показано здесь, но основные руководящие принципы этих вычислений совпадают с изложенными.

42. Разложение арктангенса и вычисление π. Если $|x| < 1$, то

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots,$$

ибо справа здесь написана геометрическая прогрессия со знаменателем $q = -x^2$. Проинтегрируем это равенство по промежутку $[0, z]$, где $0 < z < 1$. В результате мы получим:

$$\int_0^z \frac{dx}{1+x^2} = \int_0^z dx - \int_0^z x^2 dx + \int_0^z x^4 dx - \dots,$$

откуда

$$\operatorname{arctg} z = z - \frac{z^3}{3} + \frac{z^5}{5} - \frac{z^7}{7} + \dots,$$

или, возвращаясь к букве x,

$$\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots \tag{55}$$

Обе части этого равенства меняют знак вместе с x. Поэтому оно доказано для $-1 < x < 1$. При $x = 1$ это равенство принимает вид

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \dots, \tag{56}$$

причём написанный здесь ряд сходится, что видно из теоремы Лейбница. Однако предшествующими рассуждениями равенство (56) не доказано. Тем не менее, оно верно.

Для доказательства этого будем исходить не из бесконечного ряда $1-x^2+x^4-x^6+\ldots$, а из конечной суммы

$$1-x^2+x^4-x^6+\ldots+x^{2n-2}-x^{2n}=\frac{1-x^{2n+2}}{1+x^2}. \qquad (57)$$

Отсюда:

$$\frac{1}{1+x^2}=(1-x^2+x^4-x^6+\ldots+x^{2n-2}-x^{2n})+\frac{x^{2n+2}}{1+x^2}$$

и, интегрируя от 0 до 1, находим:

$$\operatorname{arctg} 1=\int_0^1\frac{dx}{1+x^2}=$$

$$=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\ldots+\frac{1}{2n-1}-\frac{1}{2n+1}+\int_0^1\frac{x^{2n+2}\,dx}{1+x^2}. \qquad (58)$$

Для интеграла в правой части получаем оценку:

$$0<\int_0^1\frac{x^{2n+2}\,dx}{1+x^2}<\int_0^1 x^{2n+2}\,dx=\frac{1}{2n+3};$$

следовательно,

$$\lim_{x\to\infty}\int_0^1\frac{x^{2n+2}\,dx}{1+x^2}=0 \qquad (59)$$

и

$$\operatorname{arctg} 1=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\ldots,$$

что и требовалось доказать.

Формула (56) представляет несомненный теоретический интерес, ибо даёт очень простое выражение числа π через натуральные числа (значительно более простое, чем формула Валлиса). Однако в качестве средства вычисления π этот ряд практически непригоден из-за его чрезвычайно медленной сходимости [1]).

Более удобное разложение π мы получим, если в основной формуле (55) положим $x=\frac{\sqrt{3}}{3}$. Это приводит к равенству

$$\frac{\pi}{6}=\frac{\sqrt{3}}{3}\left(1-\frac{1}{9}+\frac{1}{45}-\frac{1}{189}+\frac{1}{729}-\frac{1}{2673}+\ldots\right),$$

---

[1]) Чтобы получить π с двумя верными знаками после запятой, надо сложить пятьдесят (!) членов ряда (56).

откуда

$$\pi = 2\sqrt{3}\left(1 - \frac{1}{9} + \frac{1}{45} - \frac{1}{189} + \frac{1}{729} - \frac{1}{2673} + \ldots\right) \qquad (60)$$

Если ограничиться выписанными членами, то согласно замечаниям, сделанным в п° 36 к теореме Лейбница, ошибка будет положительна и меньше первого отброшенного члена, т. е. меньше, чем

$$\frac{2\sqrt{3}}{9477} < 0{,}0005,$$

так что ряд (60) даёт уже и практически приемлемое средство вычисления π. Тем не менее, для получения π с большой точностью и ряд (60) мало удобен. Для этой цели приходится [исходя из той же формулы (55)] применять некоторые искусственные приёмы. Остановимся на одном из них.

Введём в рассмотрение величину

$$\alpha = \operatorname{arctg}\frac{1}{5}$$

и пусть

$$\beta = 4\alpha - \frac{\pi}{4}.$$

Тогда

$$\pi = 16\alpha - 4\beta. \qquad (61)$$

Покажем, что, пользуясь формулой (55), мы можем вычислить величины 16α и 4β. В самом деле, полагая в (55) $x = \frac{1}{5}$ и умножая на 16, получим:

$$16\alpha = \frac{16}{5} - \frac{16}{3 \cdot 5^3} + \frac{16}{5 \cdot 5^5} - \frac{16}{7 \cdot 5^7} + \frac{16}{9 \cdot 5^9} - \frac{16}{11 \cdot 5^{11}} + \frac{16}{13 \cdot 5^{13}} - \ldots$$

Если ограничиться выписанными членами, то согласно замечаниям к теореме Лейбница ошибка Δ' будет удовлетворять неравенству

$$-\frac{1}{2 \cdot 10^{10}} < -\frac{16}{15 \cdot 5^{15}} < \Delta' < 0.$$

Складывая положительные члены, находим:

$$\begin{aligned}
\frac{16}{5} &= 3{,}20000000000 \\
\frac{16}{5 \cdot 5^5} &= 0{,}00102400000 \\
\frac{16}{9 \cdot 5^9} &= 0{,}00000091022\ (+) \\
\frac{16}{13 \cdot 5^{13}} &= 0{,}00000000101\ (-) \\
\hline
& 3{,}20102491123
\end{aligned}$$

Ошибка этого числа лежит между $-\frac{1}{2 \cdot 10^{11}}$ и $+\frac{1}{2 \cdot 10^{11}}$.

Сумма отрицательных членов такова:

$$\frac{16}{3\cdot 5^3} = 0{,}04266666667\ (-)$$
$$\frac{16}{7\cdot 5^7} = 0{,}00002925714\ (+)$$
$$\frac{16}{11\cdot 5^{11}} = 0{,}00000002979\ (-)$$
$$0{,}04269595360.$$

Ошибка этой суммы больше, чем $-\frac{1}{10^{11}}$ и меньше, чем $+\frac{1}{2\cdot 10^{11}}$.

Производя вычитание и учитывая оценки всех ошибок, находим:

$$3{,}15832895757 < 16\alpha < 3{,}15832895765. \tag{62}$$

Переходя к отысканию 4β, прежде всего заметим, что β лежит между $-\frac{\pi}{2}$ и $+\frac{\pi}{2}$. Действительно, так как $\alpha = \operatorname{arctg}\frac{1}{5}$, то $0 < \alpha < \frac{1}{5}$; стало быть, $0 < 4\alpha < \frac{4}{5}$, откуда $-\frac{\pi}{4} < 4\alpha - \frac{\pi}{4} < \frac{4}{5} - \frac{\pi}{4}$. Поэтому [1])

$$\beta = \operatorname{arctg}(\operatorname{tg}\beta).$$

С другой стороны,

$$\operatorname{tg}\beta = \operatorname{tg}\left(4\alpha - \frac{\pi}{4}\right) = \frac{\operatorname{tg} 4\alpha - 1}{1 + \operatorname{tg} 4\alpha}. \tag{63}$$

Замечая, что $\operatorname{tg}\alpha = \frac{1}{5}$, находим последовательно:

$$\operatorname{tg} 2\alpha = \frac{2\operatorname{tg}\alpha}{1 - \operatorname{tg}^2\alpha} = \frac{5}{12}, \quad \operatorname{tg} 4\alpha = \frac{2\operatorname{tg} 2\alpha}{1 - \operatorname{tg}^2 2\alpha} = \frac{120}{119}.$$

Отсюда и из (63) следует, что

$$\operatorname{tg}\beta = \frac{1}{239}.$$

Сопоставляя это с основной формулой (55), находим:

$$4\beta = \frac{4}{239} - \frac{4}{3\cdot 239^3} + \frac{4}{5\cdot 239^5} - \dots$$

Ограничиваясь всего лишь двумя (!) членами ряда, мы сделаем ошибку Δ'', удовлетворяющую неравенству

$$0 < \Delta'' < \frac{4}{5\cdot 239^5} < \frac{1}{2\cdot 10^{11}}.$$

Так как

$$\frac{4}{239} = 0{,}01673640167\ (+),$$
$$\frac{4}{3\cdot 239^3} = 0{,}00000009767\ (-),$$

то разность этих чисел равна

$$0{,}01673630400,$$

---

[1]) Напомним, что символом $\operatorname{arctg} m$ обозначается тот из углов, имеющих тангенс, равный m, *который лежит между* $-\frac{\pi}{2}$ и $+\frac{\pi}{2}$.

причём ошибка разности лежит между 0 и 10^{-11}. Сопоставляя это с оценкой для Δ'', находим:

$$0{,}01673630400 < 4\Delta < 0{,}01673630402. \tag{64}$$

Из (61), (62) и (64) получаем:

$$3{,}14159265355 < \pi < 3{,}14159265365.$$

Отсюда

$$\pi = 3{,}1415926536,$$

причём все выписанные здесь знаки верны!

43. Общие замечания по поводу разложения функций в степенные ряды. Рассмотрим какую-либо функцию $f(x)$, заданную в промежутке $[A, B]$, и пусть a — некоторая точка этого промежутка. Поставим вопрос о возможности представления $f(x)$ степенным рядом, расположенным по степеням разности $x - a$.

Чтобы не усложнять дела, будем предполагать точку a внутренней[1]) точкой $[A, B]$, т. е. считать $A < a < B$. Допустим, что для всех x, удовлетворяющих неравенству $a - r < x < a + r$, где $r > 0$, упомянутое представление возможно, т. е. что для всех этих x будет

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots \tag{65}$$

Само собой разумеется, что весь промежуток $(a - r, a + r)$ мы считаем содержащимся как в $[A, B]$, так и в промежутке сходимости написанного здесь ряда (так что радиус сходимости R этого ряда не меньше, чем r).

Полагая в (65) $x = a$, мы находим:

$$c_0 = f(a). \tag{66}$$

Так как, далее, внутри промежутка сходимости степенной ряд можно почленно дифференцировать, то для всех x из $(a - r, a + r)$ должна существовать производная $f'(x)$ и выполняться равенство

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \dots, \tag{67}$$

откуда, полагая $x = a$, мы находим:

$$c_1 = f'(a).$$

Применяя те же соображения к ряду (67), мы убеждаемся в существовании $f''(x)$ и равенстве

$$f''(x) = 2c_2 + 3 \cdot 2c_3(x-a) + 4 \cdot 3 \cdot c_4(x-a)^2 + \dots,$$

откуда

$$c_2 = \frac{f''(a)}{2!}.$$

---

[1]) В случае $a = A$ (или $a = B$) никаких принципиальных изменений в нижеприводимых рассуждениях не потребовалось бы, но пришлось бы немного изменить их редакцию.

Продолжая эти рассуждения, мы приходим к следующему выводу:

Теорема 1. *Если функция $f(x)$ в некотором промежутке $(a-r,\ a+r)$ представляется степенным рядом, расположенным по степеням $x-a$, то эта функция имеет в упомянутом промежутке производные всех порядков, а само разложение обязательно таково:*

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\\ +\frac{f'''(a)}{3!}(x-a)^3+\dots \qquad (68)$$

Написанный здесь ряд называется *рядом Тейлора* функции $f(x)$. Поэтому доказанная теорема означает, что *если функция разлагается в степенной ряд, то это — обязательно её ряд Тейлора.*

Не нужно думать, однако, что всякая функция разлагается в степенной ряд. Прежде всего может оказаться, что для функции нельзя даже составить её ряда Тейлора. Например, у функции

$$f(x)=\sqrt[3]{x}$$

при $a=0$ ряд Тейлора не существует, ибо

$$f'(x)=\frac{1}{3\sqrt[3]{x^2}},$$

а это выражение теряет смысл при $x=0$.

Далее, может оказаться, что, хотя ряд Тейлора и может быть составлен, но его радиус сходимости равен нулю, так что равенство (68) справедливо лишь для $x=a$. Практически такое равенство совершенно бесполезно.

Наконец, возможен и такой случай, когда функция имеет производные всех порядков и её ряд Тейлора сходится в некотором промежутке, но его сумма не совпадает с $f(x)$ нигде, кроме точки $x=a$.

Таким образом, даже бесконечное число раз дифференцируемые функции должны удовлетворять некоторым дополнительным условиям для того, чтобы разлагаться в степенной ряд. Не входя в полное исследование этого вопроса, ограничимся лишь одной теоремой, которая по существу была нами уже доказана в п° 12.

Теорема 2. *Пусть $f(x)$ задана в промежутке $[A,\ B]$ и имеет там производные любого порядка. Если существует такая постоянная K, что при всех натуральных n и при всех x из $[A,\ B]$ оказывается*

$$|f^{(n)}(x)|\leqslant K,$$

то равенство (68) справедливо для любой пары точек x и a из $[A,\ B]$.

В самом деле, как было показано в п° 12, в условиях теоремы остаточный член $R_n(x)$ формулы Тейлора

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_n(x)$$

стремится к нулю с возрастанием n. Но это, очевидно, совершенно равносильно равенству (68), ибо сумма

$$f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n$$

есть частичная сумма ряда Тейлора функции $f(x)$.

В частности[1]), для функций $\sin x$, $\cos x$, e^x при всех действительных x справедливы разложения

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots, \tag{69}$$

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots, \tag{70}$$

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\ldots \tag{71}$$

Формулы (69) и (70) могут быть использованы для фактического вычисления $\sin x$ и $\cos x$. Мы не будем уже останавливаться на этом вопросе, ибо основные принципы здесь те же, что и при вычислении логарифмов или арктангенсов. Обратим внимание читателя лишь на то обстоятельство, что, желая найти синус (или косинус) угла, заданного в градусах, надо сначала найти величину этого угла в радианах и именно эту величину подставлять вместо x в формулы (69) и (70).

Остановимся в заключение ещё на одном возможном применении формулы (65). Именно, если эта формула справедлива в некотором промежутке $(a-r,\ a+r)$, то, как было установлено в п° 40, для всякого замкнутого промежутка $[p,\ q]$, содержащегося в $(a-r,\ a+r)$, будет

$$\int_p^q f(x)\,dx=c_0\int_p^q dx+c_1\int_p^q (x-a)\,dx+c_2\int_p^q (x-a)^2\,dx+\ldots \tag{72}$$

Так как все интегралы, находящиеся в правой части, вычисляются без труда, то формула (72) доставляет способ приближённого вычисления интеграла

$$\int_p^q f(x)\,dx,$$

[1]) Как было бегло упомянуто в п° 39.

не требующий нахождения первообразной функции для $f(x)$ и пригодный даже в тех случаях, когда эта первообразная не выражается элементарно.

Вычислим, например, интеграл [1])

$$\int_0^1 \frac{\sin x}{x}\,dx$$

с точностью до 0,0001. На основании формулы (69) имеем:

$$\int_0^1 \frac{\sin x}{x}\,dx = \int_0^1 \left(1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots\right) dx,$$

откуда

$$\int_0^1 \frac{\sin x}{x}\,dx = 1 - \frac{1}{3\cdot 3!} + \frac{1}{5\cdot 5!} - \ldots$$

Так как этот ряд удовлетворяет условиям теоремы Лейбница, то, ограничиваясь выписанными членами, мы делаем ошибку Δ такую, что

$$0 > \Delta > -\frac{1}{7\cdot 7!} > -0{,}00003.$$

Так как

$$1 + \frac{1}{5\cdot 5!} = 1{,}00167\ (-),$$

$$\frac{1}{3\cdot 3!} = 0{,}05556\ (-),$$

то с учётом всех ошибок находим:

$$\int_0^1 \frac{\sin x}{x}\,dx = 0{,}9461,$$

причём все выписанные знаки верны.

Мы видим, что способ приближённого интегрирования при помощи рядов бывает более удобен, чем изложенный в п° 27.

44. Биномиальный ряд. Знаменитая формула, изучаемая в школе под названием «бинома Ньютона», фактически была известна ещё задолго до Ньютона. Ньютону же принадлежит заслуга её распространения на случай не натуральных показателей.

Поставим вопрос о разложении функции

$$f(x) = (1+x)^\mu$$

---

[1]) Здесь как раз $\int \frac{\sin x}{x}\,dx$ не выражается через элементарные функции.

в ряд, расположенный по степеням x. Здесь

$$f'(x)=\mu(1+x)^{\mu-1},\quad f''(x)=\mu(\mu-1)(1+x)^{\mu-2},$$
$$f'''(x)=\mu(\mu-1)(\mu-2)(1+x)^{\mu-3}$$

и вообще

$$f^{(n)}(x)=\mu(\mu-1)\dots(\mu-n+1)(1+x)^{\mu-n},$$

что можно подтвердить с помощью полной индукции.

Таким образом,

$$f(0)=1,\quad f'(0)=\mu,\quad f''(0)=\mu(\mu-1),\quad \dots$$
$$\dots,\ f^{(n)}(0)=\mu(\mu-1)\dots(\mu-n+1),\dots$$

и ряд Тейлора нашей функции таков:

$$1+\mu x+\frac{\mu(\mu-1)}{2!}x^2+\frac{\mu(\mu-1)(\mu-2)}{3!}x^3+\dots \qquad (73)$$

Этот ряд называется *биномиальным*. Если $\mu=0,\ 1,\ 2$, то ряд (73) принимает соответственно вид

$$1,\quad 1+x,\quad 1+2x+x^2,$$

т. е. превращается в конечный многочлен. Нетрудно видеть, что это явление имеет место всегда, когда μ — целое неотрицательное число. Для таких значений μ равенство функции $(1+x)^\mu$ и суммы её ряда Тейлора, т. е. равенство

$$(1+x)^\mu=1+\mu x+\frac{\mu(\mu-1)}{2!}x^2+\dots+\frac{\mu(\mu-1)\dots 1}{\mu!}x^\mu$$

справедливо при всех действительных (и даже комплексных!) x, что и составляет содержание элементарной теоремы о биноме Ньютона.

Если μ не есть целое неотрицательное число, то ряд (73) существенно бесконечен, и прежде всего встаёт вопрос о его промежутке сходимости.

Применим для решения этого вопроса признак Даламбера. У нас

$$a_n=\frac{\mu(\mu-1)\dots(\mu-n+1)}{n!}x^n,$$
$$a_{n+1}=\frac{\mu(\mu-1)\dots(\mu-n+1)(\mu-n)}{(n+1)!}x^{n+1}.$$

Стало быть,

$$\frac{a_{n+1}}{a_n}=\frac{\mu-n}{n+1}x$$

и

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=|x|.$$

Поэтому радиус сходимости ряда (73) есть 1. Мы не будем исследовать поведение ряда на концах ± 1 промежутка сходимости[1]), а ограничимся изучением его суммы в открытом промежутке $(-1, +1)$.

Выше уже указывалось, что бывают случаи, когда ряд Тейлора некоторой функции сходится, но его сумма не совпадает с этой функцией. Поэтому из того обстоятельства, что ряд (73) сходится в $(-1, +1)$, ещё не следует, что сумма его равна функции $(1+x)^\mu$. Доказать этот факт на основании теоремы 2 из п° 43 нельзя, так как условия этой теоремы здесь оказываются не выполненными. Можно было бы доказать его, пользуясь другими выражениями остаточного члена формулы Тейлора, более сложными, чем установленное в п° 12. Поскольку, однако, этих выражений мы не выводили, то и этот путь оказывается для нас закрытым. В то же время упомянутый факт весьма важен. Поэтому мы всё же докажем его, хотя и несколько искусственным способом.

Обозначим сумму ряда (73) через $S(x)$. Так как степенной ряд можно почленно дифференцировать, то при всех x из $(-1, +1)$ будет

$$S'(x) = \mu + \frac{\mu(\mu-1)}{1!}x + \frac{\mu(\mu-1)(\mu-2)}{2!}x^2 + \dots,$$

откуда

$$\frac{S'(x)}{\mu} = 1 + \frac{\mu-1}{1!}x + \frac{(\mu-1)(\mu-2)}{2!}x^2 + \dots \qquad (74)$$

Умножим равенство (74) на x и сложим полученное равенство с (74). В результате окажется

$$\frac{1+x}{\mu}S'(x) = \left\{\begin{array}{l} 1 + \frac{\mu-1}{1!}x + \frac{(\mu-1)(\mu-2)}{2!}x^2 + \frac{(\mu-1)(\mu-2)(\mu-3)}{3!}x^3 + \dots \\ \quad + x + \frac{(\mu-1)}{1!}x^2 \quad + \frac{(\mu-1)(\mu-2)}{2!}x^3 \quad + \dots, \end{array}\right.$$

откуда

$$\frac{1+x}{\mu}S'(x) = 1 + \mu x + \frac{\mu(\mu-1)}{2!}x^2 + \frac{\mu(\mu-1)(\mu-2)}{3!}x^3 + \dots$$

Правая часть последнего равенства есть не что иное, как $S(x)$. Таким образом,

$$(1+x)S'(x) = \mu S(x). \qquad (75)$$

---

[1]) Сходимость или расходимость ряда (73) в точках ± 1 зависит от значения μ.

Заметив это, рассмотрим функцию

$$\psi(x)=\frac{S(x)}{(1+x)^{\mu}},$$

заданную в $(-1, +1)$. Её производная такова:

$$\psi'(x)=\frac{S'(x)(1+x)^{\mu}-S(x)\,\mu(1+x)^{\mu-1}}{(1+x)^{2\mu}}=\frac{(1+x)S'(x)-\mu S(x)}{(1+x)^{\mu+1}}.$$

В силу (75) оказывается

$$\psi'(x)=0.$$

Значит, в $(-1, +1)$ функция $\psi(x)$ постоянна. Но непосредственно видно, что $\psi(0)=1$. Следовательно, и при всех прочих x тоже будет $\psi(x)=1$, откуда

$$S(x)=(1+x)^{\mu}.$$

Итак, доказана

Теорема 1. *При* $-1<x<1$ *будет*

$$(1+x)^{\mu}=1+\mu x+\frac{\mu(\mu-1)}{2!}x^2+\frac{\mu(\mu-1)(\mu-2)}{3!}x^3+\ldots, \quad (76)$$

каково бы ни было μ.

В частности, при $\mu=1$ формула (76) превращается в формулу суммирования геометрической прогрессии

$$\frac{1}{1+x}=1-x+x^2-x^3+\ldots$$

Отметим ещё два частных случая формулы (76). Если $\mu=-\frac{1}{2}$, то

$$\mu(\mu-1)(\mu-2)\ldots(\mu-n+1)=(-1)^n\frac{1\cdot 3\cdot 5\ldots(2n-1)}{2^n}=$$
$$=(-1)^n\frac{(2n-1)!!}{2^n}.$$

С другой стороны,

$$2^n\cdot n!=(2n)!!.$$

Значит,

$$\frac{1}{\sqrt{1+x}}=1-\frac{1!!}{2!!}x+\frac{3!!}{4!!}x^2-\frac{5!!}{6!!}x^3+\ldots \quad (77)$$

Аналогично [1])

$$\sqrt{1+x}=1+\frac{1}{2}x-\frac{1!!}{4!!}x^2+\frac{3!!}{6!!}x^3-\frac{5!!}{8!!}x^4+\ldots \quad (78)$$

---

[1]) Коэффициент при x^n в формуле (78), когда $n\geqslant 2$, таков:

$$(-1)^{n-1}\frac{(2n-3)!!}{(2n)!!}.$$

Остановимся на некоторых применениях формулы (76).

I. ИЗВЛЕЧЕНИЕ КОРНЕЙ. Пусть требуется извлечь корень степени m из натурального числа A, которое не является точной m-й степенью.

Представим A в форме

$$A = a^m + b,$$

где a — натуральное, а b — целое число, причём[1]) $|b| < a^m$. Тогда

$$\sqrt[m]{A} = a\sqrt[m]{1 + \frac{b}{a^m}}.$$

Число $x = \frac{b}{a^m}$ лежит в $(-1, +1)$ и потому функция

$$\sqrt[m]{1 + \frac{b}{a^m}} = (1 + x)^{\frac{1}{m}}$$

разлагается в ряд по степеням x, сходящийся тем быстрее, чем меньше $|x|$. Взяв достаточно большое число членов этого ряда, мы получим интересующий нас корень с любой степенью точности.

Оценку совершаемой при этом ошибки особенно удобно производить в тех случаях, когда $x > 0$. Действительно, при этом условии и при $\mu = \frac{1}{m}$, где $m > 1$, ряд

$$\mu x + \frac{\mu(\mu-1)}{2!}x^2 + \frac{\mu(\mu-1)(\mu-2)}{3!}x^3 + \ldots$$

заведомо будет знакочередующимся, и модуль его общего члена будет убывать[2]) с возрастанием n. Поэтому абсолютная величина ошибки будет меньше первого отброшенного члена ряда, а знак ошибки совпадёт со знаком этого члена.

Например, ошибка равенства

$$\sqrt[5]{38} = 2\sqrt[5]{1 + \frac{3}{16}} = 2\left(1 + \frac{3}{80} - \frac{9}{3200} + \frac{81}{256000}\right)$$

по абсолютной величине меньше, чем

$$\frac{1701}{20480000} < 0{,}0001.$$

---

[1]) Такое представление возможно бесчисленным множеством способов. Достаточно взять такое натуральное a, чтобы оказалось $a^m > A$ и положить $b = A - a^m$.

[2]) Модуль отношения двух соседних членов равен $\left|\frac{\mu - n + 1}{n + 1}x\right|$. Это — число, меньшее единицы $\left(\text{если } \mu = \frac{1}{m} \text{ и } |x| < 1\right)$.

В случае, когда $x<0$, столь просто ошибка не оценивается. Не рассматривая вопроса об оценке ошибки в общем виде, остановимся лишь на случае равенства [1])

$$\sqrt{1+x}=1+\frac{x}{2}, \tag{79}$$

получающегося из формулы (78), если в этой формуле сохранить справа только два первых члена.

В указанном случае дело сводится к оценке суммы

$$\frac{1!!}{4!!}x^2+\frac{3!!}{6!!}|x|^3+\frac{5!!}{8!!}x^4+\frac{7!!}{10!!}|x|^5+\dots$$

Так как $|x|<1$, то эта сумма не больше, чем

$$x^2\sum_{n=1}^{\infty}\frac{(2n-1)!!}{(2n+2)!!}.$$

Рассмотрим сначала дробь

$$q=\frac{(2n-1)!!}{(2n)!!}=\frac{1\cdot3\cdot5\cdot7\dots(2n-1)}{2\cdot4\cdot6\cdot8\dots(2n)}.$$

Её квадрат можно записать так:

$$q^2=\frac{1\cdot3}{2^2}\cdot\frac{3\cdot5}{4^2}\cdot\frac{5\cdot7}{6^2}\dots\frac{(2n-3)(2n-1)}{(2n-2)^2}\cdot\frac{2n-1}{2n}\cdot\frac{1}{2n}.$$

Замечая, что все дроби

$$\frac{1\cdot3}{2^2},\ \frac{3\cdot5}{4^2},\ \dots,\ \frac{(2n-3)(2n-1)}{(2n-2)^2},\ \frac{2n-1}{2n}$$

меньше единицы, находим:

$$q^2<\frac{1}{2n}.$$

Отсюда следует, что

$$\frac{(2n-1)!!}{(2n+2)!!}<\frac{1}{\sqrt{2n}}\cdot\frac{1}{2n+2}<\frac{1}{2\sqrt{2}}\cdot\frac{1}{n\sqrt{n}}.$$

Поэтому ошибка формулы (79) меньше, чем

$$\frac{x^2}{2\sqrt{2}}\sum_{n=1}^{\infty}\frac{1}{n\sqrt{n}}.$$

Чтобы оценить эту последнюю сумму, применим способ, часто употребляющийся в аналогичных случаях. Именно, рассмотрим функцию

$$\varphi(x)=\frac{1}{x\sqrt{x}}.$$

---

[1]) Оценка точности этого равенства используется ниже.

Так как эта функция убывает при $1 \leqslant x \leqslant +\infty$, то

$$\text{при } 1 \leqslant x \leqslant 2 \text{ будет } \frac{1}{2\sqrt{2}} \leqslant \frac{1}{x\sqrt{x}},$$

$$\text{при } 2 \leqslant x \leqslant 3 \text{ будет } \frac{1}{3\sqrt{3}} \leqslant \frac{1}{x\sqrt{x}},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\text{при } m-1 \leqslant x \leqslant m \text{ будет } \frac{1}{m\sqrt{m}} \leqslant \frac{1}{x\sqrt{x}}.$$

Интегрируя эти неравенства, получим:

$$\frac{1}{2\sqrt{2}} \leqslant \int_1^2 \frac{dx}{x\sqrt{x}},$$

$$\frac{1}{3\sqrt{3}} \leqslant \int_2^3 \frac{dx}{x\sqrt{x}},$$

$$\cdots\cdots\cdots\cdots$$

$$\frac{1}{m\sqrt{m}} \leqslant \int_{m-1}^{m} \frac{dx}{x\sqrt{x}}.$$

Стало быть,

$$\frac{1}{2\sqrt{2}} + \frac{1}{3\sqrt{3}} + \ldots + \frac{1}{m\sqrt{m}} \leqslant \int_1^m \frac{dx}{x\sqrt{x}} = \left[\frac{-2}{\sqrt{x}}\right]_1^m = 2 - \frac{2}{\sqrt{m}}.$$

Увеличивая m и переходя к пределу, находим:

$$\frac{1}{2\sqrt{2}} + \frac{1}{3\sqrt{3}} + \frac{1}{4\sqrt{4}} + \ldots \leqslant 2,$$

откуда

$$\sum_{n=1}^{\infty} \frac{1}{n\sqrt{n}} \leqslant 3.$$

Таким образом, доказана

Теорема 2. *Если* $|x| < 1$, *то ошибка равенства* (79) *по абсолютной величине меньше, чем* $\frac{3x^2}{2\sqrt{2}}$.

Приведём интересный пример использования этой оценки.

II. ПРАВИЛО П. Л. ЧЕБЫШЕВА ДЛЯ ПРИБЛИЖЁННОГО СПРЯМЛЕНИЯ ДУГИ ОКРУЖНОСТИ. Рассмотрим дугу $\overset{\smile}{ABC}$ окружности радиуса R, на которую опирается центральный угол $2x$, меньший 180°. Длина s этой дуги равна $2Rx$. П. Л. Чебышев предложил приближённый способ построения этой длины при помощи циркуля и линейки.

Чтобы изложить этот способ, опустим из центра O нашей окружности перпендикуляр OD на хорду AC и обозначим через B точку пересечения этого перпендикуляра с дугой (рис. 46). Отрезок BD называется *стрелкой*

дуги. Ясно, что сумма $AB+BC$ боковых сторон равнобедренного треугольника ABC (имеющего основанием хорду AC, а высотой стрелку BD) меньше s. Постараемся отодвинуть точку B вдоль перпендикуляра OD в такое положение E, чтобы точка E могла быть построена (исходя из радиуса R) при помощи циркуля и линейки и чтобы сумма $AE+EC$ боковых сторон равнобедренного треугольника AEC воспроизводила дугу s по возможности более точно (ниже мы уточняем смысл последнего требования).

Обозначим через h отношение ED к BD:

$$ED=h\cdot BD.$$

В определении величины h и заключается поставленная нами задача.

Так как

$$AD=R\sin x, \qquad BD=R(1-\cos x),$$

то, обозначая сумму $AE+EC$ через s^*, получаем:

$$s^*=2R\sqrt{\sin^2 x+h^2(1-\cos x)^2}.$$

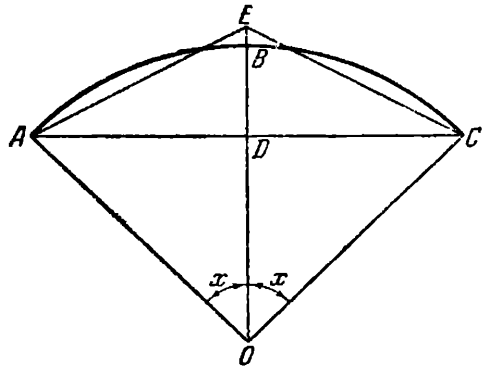


Рис. 46.

Согласно формулам (69) и (70) будет

$$\sin x=x-\frac{x^3}{6}, \quad \cos x=1-\frac{x^2}{2},$$

причём ошибки этих приближённых равенств по модулю меньше, чем $\frac{x^5}{120}$ и $\frac{x^4}{24}$. Поэтому точные равенства таковы:

$$\sin x=x-\frac{x^3}{6}+\theta_1\frac{x^5}{120}, \qquad \cos x=1-\frac{x^2}{2}+\theta_2\frac{x^4}{24},$$

где θ_1 и θ_2 по абсолютной величине меньше единицы.

В таком случае

$$\sin^2 x=x^2-\frac{x^4}{3}+x^5\left(\frac{1}{36}+\frac{\theta_1}{60}+\frac{\theta_1^2x^4}{14400}-\frac{\theta_1x^2}{360}\right),$$

$$h^2(1-\cos x)^2=h^2\frac{x^4}{4}+h^2x^6\left(\frac{\theta_2^2x^2}{576}-\frac{\theta_2}{24}\right).$$

Учитывая, что $0<x<\frac{\pi}{2}<1{,}6$, можем утверждать, что $x^2<3$, откуда

$$\left|\frac{1}{36}+\frac{\theta_1}{60}+\frac{\theta_1^2x^4}{14400}-\frac{\theta_1x^2}{360}\right|<\frac{1}{36}+\frac{1}{60}+\frac{1}{1600}+\frac{1}{120}<0{,}055,$$

$$\left|\frac{\theta_2^2x^2}{576}-\frac{\theta_2}{24}\right|<\frac{1}{192}+\frac{1}{24}<0{,}048.$$

Значит,

$$s^*=2R\sqrt{x^2+\left(\frac{h^2}{4}-\frac{1}{3}\right)x^4+\alpha x^6} \qquad (|\alpha|<0{,}055+0{,}048h^2)$$

или

$$s^*=2Rx\sqrt{1+\left(\frac{h^2}{4}-\frac{1}{3}\right)x^2+\alpha x^4}.$$

Так как истинное значение дуги s есть $2Rx$, то Чебышев считает, что наилучшим[1]) будет такое значение h, при котором под радикалом исчезнет член, содержащий x^2, т. е.

$$h=\frac{2}{\sqrt{3}}.$$

При этом выборе h точка E действительно строится циркулем и линейкой. Кроме того, при этом h оказывается

$$s^*=2Rx\sqrt{1+\alpha x^4},$$

где

$$|\alpha|<0{,}055+0{,}064=0{,}119<0{,}12.$$

Заметим теперь, что

$$x^4<(1{,}6)^4<7.$$

Значит, $|\alpha x^4|<1$ и по вышеприведённой теореме 2 ошибка равенства

$$\sqrt{1+\alpha x^4}=1+\frac{\alpha x^4}{2}$$

по модулю меньше, чем $\frac{3\alpha^2}{2\sqrt{2}}x^8$, и, тем более, меньше, чем $0{,}02x^8$. Таким образом,

$$\sqrt{1+\alpha x^4}=1+\frac{\alpha x^4}{2}+\beta x^8,$$

где $|\beta|<0{,}02$. Умножая на $2Rx$, находим:

$$s^*=2Rx+\alpha Rx^5+2\beta Rx^9.$$

Заметим, наконец, что

$$|\alpha x^5+2\beta x^9|<x^5(0{,}12+0{,}04x^4)<0{,}4x^5,$$

откуда

$$s^*=s+\rho,$$

где $|\rho|<0{,}4Rx^5$.

Резюмируя, приходим к следующему результату:

Теорема 3 (П. Л. Чебышев). *Дуга окружности приблизительно равна сумме боковых сторон равнобедренного треугольника, у которого основанием служит хорда, стягивающая дугу, а высота равна* $\frac{2}{\sqrt{3}}$ *стрелки. Абсолютная величина ошибки этого приближённого равенства меньше, чем*

$$0{,}4Rx^5,$$

где x — половина центрального угла, опирающегося на дугу, а R — радиус окружности.

Ясно, что правило Чебышева обладает тем лучшей точностью, чем меньше x. Например, если центральный угол равен 30°, т. е.

$$x=\frac{\pi}{12},$$

---

[1]) Это и есть то специальное соглашение, которое вносит совершенно определённый и чёткий смысл в высказанное раньше несколько расплывчатое требование, чтобы сумма $AE+EC$ воспроизводила дугу s «по возможности более точно».

то ошибка будет меньше, чем $0{,}0006R$. Действительно, используя найденное в п° 42 значение π, получаем:

$$s = 0{,}523599R.$$

Правило же Чебышева даёт

$$s = 2R\sqrt{\sin^2 15^\circ + \frac{4}{3}(1-\cos 15^\circ)^2},$$

откуда

$$s = 0{,}523585R,$$

т. е. ошибка, равная $0{,}000014R$, попадает в указанные границы.

III. РАЗЛОЖЕНИЕ АРКСИНУСА. Рассмотрим в заключение вопрос о разложении функции $\arcsin x$ в ряд по степеням x. Непосредственное составление ряда Тейлора здесь было бы затруднительно, ввиду весьма громоздких выражений последовательных производных $\arcsin x$. Использование биномиального ряда позволяет обойти эту трудность. Именно, как мы видели выше,

$$\frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{(2n-1)!!}{(2n)!!} x^n.$$

Заменим здесь x на $-z$, а затем z на x^2, что даёт

$$\frac{1}{\sqrt{1-x^2}} = 1 + \sum_{n=1}^{\infty} \frac{(2n-1)!!}{(2n)!!} x^{2n}.$$

Интегрируя это равенство по промежутку $[0, x]$, где $-1 < x < 1$, мы получим искомое разложение:

$$\arcsin x = x + \frac{1!!}{2!!}\frac{x^3}{3} + \frac{3!!}{4!!}\frac{x^5}{5} + \frac{5!!}{6!!}\frac{x^7}{7} + \ldots$$

45. Очерк аналитической теории тригонометрических функций. В математическом анализе большое значение [1]) имеют тригонометрические функции, которые, однако, вводятся в математику на основании геометрических построений, совершенно чуждых анализу. Возникает вопрос, можно ли построить теорию этих функций, не прибегая к геометрическим соображениям, а оставаясь на чисто аналитической почве. Вопрос этот следует рассмотреть, потому что наряду с евклидовой существуют и другие геометрии, в связи с чем естественно возникает опасение, что результаты анализа зависят от выбора геометрии. Может быть, приняв геометрию Лобачевского [2]), мы должны будем изменить и ряд теорем анализа?

---

[1]) Напомним хотя бы тригонометрические подстановки в интегральном исчислении (п° 18, пример интеграла $\int \sqrt{1-x^2}\,dx$).

[2]) Как известно, в геометрии Лобачевского нет подобных фигур. В то же время теория подобия лежит в основе обычного построения тригонометрии.

На самом деле результаты анализа от выбора той или иной геометрии не зависят, а тригонометрические функции можно определить и изучить, совершенно не используя никаких геометрических соображений. Здесь мы имеем в виду вкратце показать, как это можно сделать.

Рассмотрим два степенных ряда

$$1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\frac{x^8}{8!}-\dots, \tag{80}$$

$$\frac{x}{1!}-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\frac{x^9}{9!}-\dots \tag{81}$$

Пользуясь признаком Даламбера, легко показать, что каждый из этих рядов сходится при всех действительных x. Обозначим суммы этих рядов соответственно через [1]) $C(x)$ и $S(x)$ и назовём их **косинусом и синусом** аргумента x. Таким образом,

$$C(x)=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\dots, \quad S(x)=\frac{x}{1!}-\frac{x^3}{3!}+\frac{x^5}{5!}-\dots$$

Из этих формул сразу видно, что

$$C(-x)=C(x), \quad S(-x)=-S(x), \tag{82}$$

т. е. что *косинус — функция чётная, а синус — нечётная.*

Далее, пользуясь теоремой о почленном дифференцировании степенных рядов, получаем:

$$C'(x)=-S(x), \qquad S'(x)=C(x). \tag{83}$$

Отсюда

$$\begin{aligned} C''(x)&=-C(x), & S''(x)&=-S(x),\\ C'''(x)&=S(x), & S'''(x)&=-C(x),\\ C^{(4)}(x)&=C(x), & S^{(4)}(x)&=S(x) \end{aligned}$$

и вообще при любом натуральном k

$$\left.\begin{aligned} C^{(4k)}(x)&=C(x), & S^{(4k)}(x)&=S(x),\\ C^{(4k+1)}(x)&=-S(x), & S^{(4k+1)}(x)&=C(x),\\ C^{(4k+2)}(x)&=-C(x), & S^{(4k+2)}(x)&=-S(x),\\ C^{(4k+3)}(x)&=S(x), & S^{(4k+3)}(x)&=-C(x). \end{aligned}\right\} \tag{84}$$

Установим теперь основную во всей теории **теорему сложения для косинуса**. Для этого отметим, что какой бы конечный

---

[1]) Как мы знаем, эти суммы суть $\cos x$ и $\sin x$. Поэтому мы могли бы и обозначить их этими символами. Однако мы предпочитаем обозначения $C(x)$ и $S(x)$, чтобы незаметно не использовать каких-либо привычных, но ещё не обоснованных аналитически, свойств этих функций.

промежуток $[A, B]$ ни взять, для него найдётся такая постоянная K, что при всех натуральных n и при всех x из $[A, B]$ будет

$$|C^{(n)}(x)| \leqslant K.$$

Действительно, в зависимости от чётности или нечётности числа n будет или $|C^{(n)}(x)| = |C(x)|$, или $|C^{(n)}(x)| = |S(x)|$, а каждая из непрерывных (см. п° 40, теорема 1) функций $C(x)$ и $S(x)$ ограничена на $[A, B]$.

Отсюда (по теореме 2 из п° 43) следует, что для любых двух точек x и a из $[A, B]$ будет

$$C(x) = C(a) + \frac{C'(a)}{1!}(x-a) + \frac{C''(a)}{2!}(x-a)^2 + \ldots \qquad (85)$$

Но ведь для любых двух точек x и a всегда можно построить содержащий их промежуток $[A, B]$. Поэтому равенство (85) имеет место вообще для любых x и a.

Заменим в (85) букву x на $x+a$. Это даёт

$$C(x+a) = C(a) + \frac{C'(a)}{1!}x + \frac{C''(a)}{2!}x^2 + \frac{C'''(a)}{3!}x^3 + \ldots$$

Если использовать соотношения (84), то последнему равенству можно придать вид

$$C(x+a) = C(a) - S(a)\frac{x}{1!} - C(a)\frac{x^2}{2!} + S(a)\frac{x^3}{3!} + C(a)\frac{x^4}{4!} - \ldots$$

Иначе говоря, коэффициентами при $\frac{x^n}{n!}$ будут только числа $C(a)$ и $S(a)$, или эти же числа, но с обратными знаками.

Объединяя попарно члены последнего ряда, получим:

$$C(x+a) = \left[C(a) - S(a)\frac{x}{1!}\right] + \left[-C(a)\frac{x^2}{2!} + S(a)\frac{x^3}{3!}\right] + \left[C(a)\frac{x^4}{4!} - S(a)\frac{x^5}{5!}\right] + \ldots$$

Так как сходящиеся ряды можно почленно вычитать, то в точности к тому же ряду мы пришли бы, вычитая почленно из ряда

$$C(a) - C(a)\frac{x^2}{2!} + C(a)\frac{x^4}{4!} - \ldots$$

ряд

$$S(a)\frac{x}{1!} - S(a)\frac{x^3}{3!} + S(a)\frac{x^5}{5!} - \ldots$$

Поэтому, вынося в последних двух рядах $C(a)$ и $S(a)$ за скобки, мы находим:

$$C(x+a) = C(a)\left[1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\right] - S(a)\left[\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots\right].$$

Ряды, стоящие здесь в скобках, суть в точности ряды (80) и (81). Стало быть,

$$C(x+a)=C(a)\,C(x)-S(a)\,S(x). \tag{86}$$

Это и есть теорема сложения для косинуса. Заменяя в этой формуле a на $-a$ и применяя формулы (82), получаем теорему вычитания

$$C(x-a)=C(a)\,C(x)+S(a)\,S(x). \tag{87}$$

Здесь a и x — любые числа. Полагая, в частности, $a=x$, и замечая, что из самого определения $C(x)$ следует равенство

$$C(0)=1, \tag{88}$$

получим известную формулу

$$C^2(x)+S^2(x)=1, \tag{89}$$

которая при обычном построении теории выводится из теоремы Пифагора. Из (89) видно, что всегда

$$|C(x)|\leqslant 1, \quad |S(x)|\leqslant 1.$$

Если продифференцировать тождества (86) и (87) и применить формулы (83), то мы придём к теоремам сложения и вычитания для синуса:

$$S(x+a)=S(x)\,C(a)+S(a)\,C(x), \tag{90}$$

$$S(x-a)=S(x)\,C(a)-S(a)\,C(x). \tag{91}$$

Из формул сложения получаются формулы удвоения:

$$C(2x)=C^2(x)-S^2(x),$$

$$S(2x)=2S(x)\,C(x).$$

Подобным же образом можно было бы получить ещё ряд формальных соотношений, например выражения для $S(3x)$ и т. д.

Труднее устанавливается периодичность функций $C(x)$ и $S(x)$ и связанные с ней формулы приведения.

Покажем прежде всего, что при $0<x\leqslant 2$ будет

$$S(x)>0.$$

В самом деле, ведь

$$S(x)=\frac{x}{1!}\left(1-\frac{x^2}{2\cdot 3}\right)+\frac{x^5}{5!}\left(1-\frac{x^2}{6\cdot 7}\right)+\frac{x^9}{9!}\left(1-\frac{x^2}{10\cdot 11}\right)+\ldots$$

и все скобки здесь (при $0\leqslant x\leqslant 2$) положительны.

Вспоминая, что $C'(x)=-S(x)$, мы видим, что функция $C(x)$ на промежутке [0, 2] строго убывает.

С другой стороны, записывая $C(2)$ в форме

$$C(2)=1-\frac{2^2}{2!}+\frac{2^4}{4!}-\frac{2^6}{6!}\left(1-\frac{2^2}{7\cdot 8}\right)-\frac{2^{10}}{10!}\left(1-\frac{2^2}{11\cdot 12}\right)-\ldots$$

и замечая, что все скобки здесь положительны и что

$$1-\frac{2^2}{2!}+\frac{2^4}{4!}=-\frac{1}{3},$$

мы видим, что $C(2)<0$. Отсюда и из (88) вытекает, что непрерывная функция $C(x)$ в открытом промежутке (0, 2) имеет один и только один корень. Обозначим его через [1]) $\frac{\tau}{2}$:

$$C\left(\frac{\tau}{2}\right)=0 \qquad \left(0<\frac{\tau}{2}<2\right). \tag{92}$$

В силу (89) и (92) вытекает, что

$$S\left(\frac{\tau}{2}\right)=+1, \tag{93}$$

ибо в (0, 2] функция $S(x)$ положительна.

Установив соотношения (92) и (93), получим из (86) и (90)

$$C\left(x+\frac{\tau}{2}\right)=-S(x), \quad S\left(x+\frac{\tau}{2}\right)=C(x). \tag{94}$$

В таком случае

$$\begin{aligned} C(x+\tau)&=-S\left(x+\frac{\tau}{2}\right)=-C(x),\\ S(x+\tau)&=C\left(x+\frac{\tau}{2}\right)=-S(x), \end{aligned} \tag{95}$$

и стало быть,

$$C(x+2\tau)=C(x), \quad S(x+2\tau)=S(x). \tag{96}$$

Иными словами, *функции* $C(x)$ *и* $S(x)$ *имеют период* 2τ. Попутно нами установлены и некоторые из формул приведения [например, (94) и (95)]. Ещё одна формула приведения следует из формулы (87), ибо, заменяя в этой последней формуле x через $\frac{\tau}{2}$, а a — через x, сразу находим:

$$C\left(\frac{\tau}{2}-x\right)=S(x). \tag{97}$$

В частности, отсюда и из (89) вытекает, что

$$C\left(\frac{\tau}{4}\right)=S\left(\frac{\tau}{4}\right)=\frac{\sqrt{2}}{2}. \tag{98}$$

---

[1]) Введённое здесь число τ есть не что иное, как π. Но мы, как и ныше, избегаем привычных обозначений, чтобы с полной отчётливостью установить независимость нашего изложения от геометрических соображений.

При помощи доказанных формул легко показать, что корнями функции $S(x)$ служат все точки вида $n\tau$ ($n=0, \pm 1, \pm 2, \ldots$) и только эти точки, а корнями $C(x)$ являются точки $n\tau+\frac{\tau}{2}$ ($n=0, \pm 1, \pm 2, \ldots$) и только они.

Построив, таким образом, теорию синуса и косинуса, мы можем ввести функцию «тангенс», положив

$$T(x)=\frac{S(x)}{C(x)}.$$

Это — нечётная функция, непрерывная в открытом промежутке $\left(-\frac{\tau}{2}, +\frac{\tau}{2}\right)$, строго возрастающая там и стремящаяся соответственно к $-\infty$ или к $+\infty$, когда x, оставаясь в указанном промежутке, приближается к $-\frac{\tau}{2}$ или к $+\frac{\tau}{2}$. Так как

$$T(x+\tau)=\frac{S(x+\tau)}{C(x+\tau)}=\frac{-S(x)}{-C(x)}=\frac{S(x)}{C(x)}=T(x),$$

то *периодом функции* $T(x)$ *служит* τ.

Не останавливаясь на других свойствах функции $T(x)$, отметим лишь равенства

$$T\left(\frac{\tau}{4}\right)=1, \quad T'(x)=\frac{1}{C^2(x)},$$

первое из которых следует из (98), а второе доказывается на основании правила дифференцирования частного и формул (82) и (89).

Рассматривая функцию $y=T(x)$ лишь в открытом промежутке $\left(-\frac{\tau}{2}, +\frac{\tau}{2}\right)$, где она непрерывна и строго возрастает, мы можем ввести для неё обратную функцию

$$x=a(y),$$

заданную в промежутке $-\infty<y<+\infty$ и имеющую значения, попадающие в $\left(-\frac{\tau}{2}, +\frac{\tau}{2}\right)$. Эта функция — также непрерывная и строго возрастающая. По теореме о дифференцировании обратной функции (см. п° 5) у функции $a(y)$ всюду существует производная

$$a'(y)=\frac{1}{T'(x)},$$

где $T(x)=y$. Пользуясь формулой для $T'(x)$, находим:

$$a'(y)=C^2(x),$$

откуда

$$a'(y)=\frac{C^2(x)}{C^2(x)+S^2(x)}=\frac{1}{1+\frac{S^2(x)}{C^2(x)}}=\frac{1}{1+T^2(x)}=\frac{1}{1+y^2}.$$

Возвращаясь к обычному обозначению независимой переменной через x, получаем:

$$a'(x)=\frac{1}{1+x^2}.$$

Отсюда следует, что

$$\int_0^1 \frac{dx}{1+x^2}=[a(x)]_0^1=a(1)-a(0)=\frac{\tau}{4}. \tag{99}$$

Значит, число τ можно найти с любой степенью точности, вычисляя интеграл

$$\int_0^1 \frac{dx}{1+x^2}.$$

В п° 27 мы как раз и вычисляли этот интеграл. Пользуясь найденными там результатами, можем утверждать, что с ошибкой, абсолютно меньшей 0,005, будет

$$\tau=3{,}14.$$

Другой способ вычисления τ состоит в рассмотрении функции, обратной для $S(x)$, которая существует, если изменять x в $\left[-\frac{\tau}{2}, +\frac{\tau}{2}\right]$. С помощью этой функции легко показать, что

$$\int_{-\frac{\sqrt{2}}{2}}^{+\frac{\sqrt{2}}{2}} \frac{dx}{\sqrt{1-x^2}}=\frac{\tau}{2}. \tag{100}$$

Мы не будем останавливаться на подробном доказательстве равенства (100), ибо здесь нет ничего нового по сравнению с уже рассмотренным выводом формулы (99).

Остановимся в заключение на вопросе о отождествлении рассмотренных нами функций $S(x)$, $C(x)$, $T(x)$, $a(x)$ с обычными тригонометрическими функциями $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{arctg} x$ и числа τ с числом π.

Этот вопрос можно ставить двояко. Во-первых, можно считать, что обычная теория тригонометрических функций, основанная на геометрических представлениях, уже построена. Тогда можно пользоваться всеми свойствами этих функций, устанавливаемыми в обычной теории. В частности, можно считать известными разложения $\cos x$ и $\sin x$ в степенные ряды и достаточно лишь подчеркнуть, что это суть те самые ряды (80) и (81), при помощи которых определялись

функции $C(x)$ и $S(x)$. Что касается равенства $\tau = \pi$, то оно при указанном подходе вытекает хотя бы из того, что

$$\frac{\tau}{4} = \int_0^1 \frac{dx}{1+x^2}, \quad \frac{\pi}{4} = \int_0^1 \frac{dx}{1+x^2}.$$

Более интересен другой подход, при котором мы считаем известными изложенные в настоящем п° свойства функций $C(x)$, $S(x)$ и т. д. и ставим вопрос о связи этих функций с окружностью, но обычную теорию тригонометрических функций не предполагаем известной.

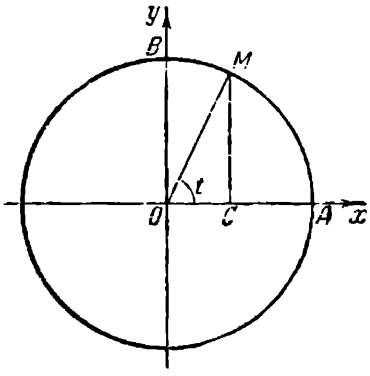


Рис. 47.

При этом подходе мы начинаем с равенства

$$C^2(t) + S^2(t) = 1,$$

которое показывает, что точка с координатами $(C(t), S(t))$ лежит на окружности $x^2 + y^2 = 1$. Отсюда сразу вытекает, что $C(t)$ и $S(t)$ соответственно равны отрезкам OC и CM (рис. 47). Это, однако, ещё не даёт нам права отождествлять их с теми отрезками, которые в обычной теории обозначаются через $\cos t$ и $\sin t$, ибо мы не знаем, где на рис. 47 изображён аргумент t.

Чтобы разобраться в этом вопросе и попутно установить геометрическое значение постоянной τ, рассмотрим верхнюю полуокружность единичной окружности. Её уравнение таково:

$$y = \sqrt{1-x^2}.$$

Значит, согласно формуле (4) из главы III длина той дуги этой полуокружности, которая расположена над отрезком $\left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right]$ оси Ox, равна

$$\int_{-\frac{\sqrt{2}}{2}}^{+\frac{\sqrt{2}}{2}} \frac{dx}{\sqrt{1-x^2}},$$

т. е. [см. (100)] равна $\frac{\tau}{2}$. В силу симметрии длина всей окружности (единичного радиуса) равна 2τ. А так как длины окружностей относятся, как диаметры, то τ оказывается не чем иным, как отношением длины окружности к её диаметру, т. е. мы отождествляем τ с π. Впредь мы и будем писать π вместо τ.

Если координаты точки M суть $C(t)$ и $S(t)$, то длина дуги BM будет [1])

$$s_{BM} = \int_0^{C(t)} \frac{dx}{\sqrt{1-x^2}} = \int_0^{S\left(\frac{\pi}{2}-t\right)} \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2} - t.$$

В таком случае длина дуги AM будет $\frac{\pi}{2} - \left(\frac{\pi}{2} - t\right) = t$, а тогда и величина угла AOM в радианах равна t.

Таким образом, чтобы построить точку $M(C(t), S(t))$, мы должны провести луч OM, наклонённый под углом t к положительной оси Ox, и найти его пересечение с единичной окружностью. Тем самым мы и убеждаемся (уже при новом подходе) в том, что $C(t)$ и $S(t)$ это суть обычные $\cos t$ и $\sin t$.

---

[1]) Чтобы получить этот результат, надо ввести функцию $x = \alpha(y)$, обратную для $y = S(x)$ $\left(\text{предполагая} -\frac{\pi}{2} \leqslant x \leqslant \frac{\pi}{2}\right)$. Тогда $\alpha'(x) = \frac{1}{\sqrt{1-x^2}}$ и

$$\int_0^{S\left(\frac{\pi}{2}-t\right)} \frac{dx}{\sqrt{1-x^2}} = \Big[\alpha(x)\Big]_0^{S\left(\frac{\pi}{2}-t\right)} = \frac{\pi}{2} - t.$$

В. Л. ГОНЧАРОВ

ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ КОМПЛЕКСНОГО ПЕРЕМЕННОГО

§ 1. Рациональные функции

Комплексное число (независимое переменное в комплексной области) обозначают обыкновенно буквой z; его действительную часть — буквой x, его мнимую часть — буквой y. Таким образом, мы имеем:

$$z = x + iy.$$

Если нужно различать несколько комплексных чисел, то добавляют значки или пользуются соответствующими буквами

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2, \ \ldots, \ \zeta = \xi + i\eta \text{ и т. п.}$$

Из курса школьной алгебры известны правила четырёх арифметических действий над комплексными числами; их можно резюмировать краткой формулой: *сложение, вычитание, умножение и деление комплексных чисел совершаются по обычным правилам алгебры, с условием заменять выражение i^2 через -1 всякий раз, как такое выражение встречается.* Например,

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2),$$
$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Остаются в силе основные законы:

$$z_1 + z_2 = z_2 + z_1, \quad z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3,$$
$$z_1 z_2 = z_2 z_1, \qquad z_1 (z_2 z_3) = (z_1 z_2) z_3,$$
$$z_1 (z_2 + z_3) = z_1 z_2 + z_1 z_3.$$

Обратные действия — вычитание и деление — выполнимы и однозначны, кроме деления на нуль; последнее невозможно. Практически деление выполняется всегда с помощью приёма умножения числителя (делимого) и знаменателя (делителя) на число, сопряжённое со знаменателем:

$$\frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + i\frac{-x_1 y_2 + x_2 y_1}{x_2^2 + y_2^2}.$$

Из отмеченных обстоятельств вытекает в качестве следствия, что *все правила рациональных алгебраических операций в комплексной области остаются те же, что и в действительной.* Так, «разность

квадратов» неизменно равна «произведению суммы на разность», сохраняется в силе формула «бинома Ньютона» (при целом положительном показателе) и т. п.

Из сказанного выше можно также заключить, что в комплексной области не теряет смысла понятие рациональной функции.

Условимся обозначать зависимое переменное, или функцию (также комплексное число), буквой w; действительную и мнимую части w — соответственно через u и v. Итак,

$$\left.\begin{array}{c} w=f(z), \\ z=x+iy, \quad w=u+iv. \end{array}\right\} \tag{1}$$

Если задана некоторая рациональная функция $w=f(z)$ комплексного переменного z, т. е. указана совокупность рациональных действий, которые надо совершить в определённом порядке над переменным z и данными постоянными (вообще говоря, также комплексными числами), чтобы получить w, то, каково бы ни было данное значение z, при условии, что не придётся делить на нуль, — значение w может быть вычислено посредством приведённых выше общеизвестных приёмов. Пусть, например,

$$w=f(z)=\frac{z}{z-1};$$

тогда, положив хотя бы

$$z=2+3i,$$

мы получим:

$$w=\frac{2+3i}{(2+3i)-1}=\frac{2+3i}{1+3i}=\frac{(2+3i)(1-3i)}{(1+3i)(1-3i)}=\frac{11}{10}-\frac{3}{10}i.$$

Можно также подставлять в формулу ряд заданных значений z и записывать результаты в табличной форме:

z	w
$2+3i$	$\frac{11-3i}{10}$
$1+i$	$1-i$
$3-2i$	$\frac{5+i}{4}$
i	$\frac{1-i}{2}$
$\frac{3-i}{2}$	$2+i$
. . .	. . .

Если в общем случае произвольной рациональной функции оставить x и y буквенными, но выполнить все действия, указанные символом f, то легко понять, что в результате мы сможем «разделить»

действительную и мнимую части в данном уравнении [1]), и окажется, что величины u и v будут некоторыми рациональными же функциями от переменных x и y

$$\left.\begin{aligned} u &= u(x, y), \\ v &= v(x, y). \end{aligned}\right\} \tag{2}$$

Так, в нашем примере, полагая $z = x + iy$ и $w = u + iv$, мы получаем:

$$u + iv = \frac{x + iy}{(x + iy) - 1},$$

и после преобразований

$$u + iv = \frac{(x^2 + y^2 - x) - iy}{(x-1)^2 + y^2},$$

так что

$$u = \frac{x^2 + y^2 - x}{(x-1)^2 + y^2}, \qquad v = -\frac{y}{(x-1)^2 + y^2}.$$

Имея такие формулы, мы могли бы вместо «$z \to w$» таблицы составить «$(x, y) \to (u, v)$» таблицу:

x	y	u	v
2	3	$\frac{11}{10}$	$-\frac{3}{10}$
1	1	1	-1
3	-2	$\frac{5}{4}$	$\frac{1}{4}$
0	1	$\frac{1}{2}$	$-\frac{1}{2}$
$\frac{3}{2}$	$-\frac{1}{2}$	2	1
...	...	...	...

Заметим, что в рассмотренном примере нельзя было бы положить z равным 1, т. е. x равным 1, а y равным 0.

Рекомендуем читателю разделить действительную и мнимую части рациональной функции в следующих примерах:

1) $w = z^2$ $\quad (u = x^2 - y^2, \quad v = 2xy)$;

2) $w = z^3$ $\quad (u = x^3 - 3xy^2, \quad v = 3x^2y - y^3)$;

3) $w = \frac{1}{z}$ $\quad \left(u = \frac{x}{x^2 + y^2}, \quad v = -\frac{y}{x^2 + y^2}\right)$;

4) $w = \frac{z - i}{z + i}$ $\quad \left(u = \frac{x^2 + y^2 - 1}{x^2 + (y+1)^2}, \quad v = -\frac{2x}{x^2 + (y+1)^2}\right)$.

[1]) Два комплексных числа равны в том и только в том случае, если в отдельности равны между собой их действительные части, а также равны между собой их мнимые части («одно комплексное равенство равносильно двум действительным»).

Геометрическое истолкование комплексных чисел как точек в «комплексной» плоскости мы предполагаем известным. Говорят об «x,y-плоскости», или о «z-плоскости».

Функциональное соотношение (1) или (2) в таком случае представляет геометрически «отображение» x,y-плоскости на u,v-плоскость, или z-плоскости на w-плоскость (см. стр. 261).

Отметим, что поскольку не определены ещё никакие операции, кроме четырёх арифметических, над комплексными числами, ни о каких функциях комплексного переменного [1]), кроме рациональных, не может быть речи. Такие символы, как, например,

$$2^i,\ \cos i,\ \lg(3i+2) \text{ и т. п.},$$

лишены для нас всякого смысла до тех пор, пока не будут даны определения показательной функции, тригонометрических функций, логарифма и т. д. в комплексной области [2]).

Задачей настоящей статьи в первую очередь и является — распространить определения элементарных функций на комплексную область.

Такого рода определения будут даны в ближайших параграфах (§§ 3—8); однако принципы, лежащие в основе этих определений и обусловливающие их логическую необходимость, будут выяснены позднее (см. §§ 9—16).

Интересующие нас определения проще всего даются с помощью рядов; поэтому нам предварительно придётся остановиться на понятии предела и понятии ряда в комплексной области.

§ 2. Пределы. Ряды

Определение *предела последовательности* [3]) чисел

$$\{z_n\} \equiv z_1,\ z_2,\ z_3,\ \ldots$$

в комплексной области — буквально то же, что и в действительной: говорят, что последовательность $\{z_n\}$ *стремится к пределу* c, или *имеет предел* c

$$\lim z_n = c, \quad z_n \to c,$$

если, как бы мало ни было положительное число ε, при достаточно больших значениях n (при $n > N_\varepsilon$) имеет место неравенство

$$|z_n - c| < \varepsilon. \tag{3}$$

---

[1]) Термин «функция» употребляется здесь в оперативном смысле, а не в смысле «соответствия».

[2]) В курсе элементарной алгебры иногда, кроме рациональных операций, вводится также операция извлечения корня целой положительной степени

$$f(z) = \sqrt[n]{z}.$$

К этому вопросу мы обратимся ниже в § 7.

[3]) О пределе функции сказано в § 9.

В левой части этого неравенства стоит знак модуля, являющийся обобщением знака абсолютной величины [1]). Нетрудно уяснить себе смысл этого неравенства: полагая

$$z_n = x_n + iy_n, \quad c = a + ib,$$

мы убеждаемся, что модуль разности

$$|z_n - c| = \sqrt{(x_n - a)^2 + (y_n - b)^2}$$

представляет собой расстояние между точками z_n и c в комплексной плоскости z; поэтому соотношение (3) может быть истолковано геометрически следующим образом: как бы ни был мал радиус ε круга с центром в точке c («ε-круг», ε-окрестность), все точки последовательности $\{z_n\}$, начиная с некоторой, в него попадают.

Весьма употребителен также знак бесконечного предела (∞ без знаков $+$ или $-$): соотношение

$$\lim z_n = \infty, \quad z_n \to \infty$$

означает, что как бы велико ни было положительное число M, при достаточно больших значениях n (при $n > N_M$) имеет место неравенство

$$|z_n| > M. \tag{4}$$

Это означает, что, как бы ни был велик радиус M круга с центром в точке O (а, впрочем, безразлично, с каким центром), все точки последовательности $\{z_n\}$, начиная с некоторой, оказываются вне его.

По отношению к конечным пределам в комплексной области справедливы те же теоремы, что и в действительной:

Если $z_n' \to c'$, $z_n'' \to c''$, то

I. $z_n' + z_n'' \to c' + c''$,

II. $z_n' - z_n'' \to c' - c''$,

III. $z_n' z_n'' \to c' c''$,

IV. $\dfrac{z_n'}{z_n''} \to \dfrac{c'}{c''}$ (при условии, что $c'' \neq 0$).

Они вытекают из того, что знак модуля обладает свойствами знака абсолютной величины (обобщая его):

$$|z' + z''| \leqslant |z'| + |z''|,$$
$$|z' z''| = |z'| \cdot |z''|.$$

---

[1]) *Модулем* z называется расстояние точки z от начала O:

$$|z| = |x + iy| = \sqrt{x^2 + y^2} \quad (z \geqslant 0).$$

См. А. П. Киселёв, Алгебра, ч. II, § 140.

Иногда приходится также ссылаться на следующие, с геометрической точки зрения, впрочем, достаточно очевидные, предложения:

a_1) Положим попрежнему

$$z_n = x_n + iy_n, \quad c = a + ib.$$

Тогда если $x_n \to a$ и $y_n \to b$, то $z_n \to c$.

a_2) Обратно, если $z_n \to c$, то $x_n \to a$ и $y_n \to b$. Доказательство следует из неравенств

$$\left.\begin{matrix} |x_n - a| \\ |y_n - b| \end{matrix}\right\} \leqslant |z_n - c| \leqslant |x_n - a| + |y_n - b|$$

(каждый из катетов прямоугольного треугольника меньше гипотенузы, а гипотенуза меньше суммы катетов).

b_1) Положим

$$z_n = r_n(\cos\theta_n + i\sin\theta_n), \quad c = \rho(\cos\omega + i\sin\omega).$$

Тогда если $\theta_n \to \omega$ и $r_n \to \rho$, то $z_n \to c$ ($\rho \neq 0$)[1]).

В самом деле, из сделанных допущений вытекает:

$$x_n = r_n\cos\theta_n \to \rho\cos\omega = a,$$
$$y_n = r_n\sin\theta_n \to \rho\sin\omega = b,$$

и остаётся сослаться на теорему a_1).

b_2) Если

$$z_n \to c \qquad (c \neq 0),$$

то при тех же обозначениях и при допущении $\rho \neq 0$

$$r_n \to \rho$$

и при надлежащем выборе аргументов θ_n[2])

$$\theta_n \to \omega.$$

В самом деле, из соотношения

$$|r_n - \rho| = \big||z_n| - |c|\big| \leqslant |z_n - c|$$

вытекает, что $r_n \to \rho$; далее, на основании a_2) получаем:

$$r_n\cos\theta_n \to \rho\cos\omega,$$
$$r_n\sin\theta_n \to \rho\sin\omega,$$

так что

$$\cos\theta_n \to \cos\omega,$$
$$\sin\theta_n \to \sin\omega$$

и отсюда (в соответствии с выбором аргументов) следует:

$$\theta_n \to \omega.$$

---

[1]) Если $r_n \to 0$, то (без какого бы то ни было условия, налагаемого на θ_n) $z_n \to 0$, и обратно.

[2]) Каждое комплексное число, отличное от нуля, имеет бесчисленное множество аргументов, определённых с точностью до слагаемых, кратных 2π. Можно принять, что аргументы чисел z_n и c фиксируются по следующему правилу: если $c > 0$, то полагаем

$$\omega = 0, \quad -\pi < \theta_n \leqslant +\pi;$$

во всех прочих случаях полагаем

$$0 < \omega < 2\pi, \quad 0 \leqslant \theta_n < 2\pi.$$

Поскольку понятие ряда приводится к понятию предела, о сходимости рядов в комплексной области можно сказать то же, что и в действительной: *ряд*

$$\sum_{n=1}^{\infty} z_n = z_1 + z_2 + \ldots + z_n + \ldots$$

сходится и имеет сумму s («сходится к сумме s»), *если, полагая*

$$s_n = z_1 + z_2 + \ldots + z_n \qquad (n = 1, 2, \ldots),$$

мы будем иметь соотношение

$$\lim s_n = s. \tag{5}$$

В теории рядов с комплексными членами существенно понятие *абсолютной сходимости*: по определению *ряд* $\sum_{n=1}^{\infty} z_n$ *сходится абсолютно, если сходится не только он сам, но и ряд, составленный из модулей его членов,* $\sum_{n=1}^{\infty} |z_n|$.

Пусть $z_n = x_n + iy_n$ $(n = 1, 2, \ldots)$; *если сходится абсолютно ряд* $\sum_{n=1}^{\infty} z_n$, *то сходятся абсолютно также ряды* $\sum_{n=1}^{\infty} x_n$ *и* $\sum_{n=1}^{\infty} y_n$, *и обратно.*

Справедливость этой теоремы следует из того, что $|x_n| \leqslant |z_n|$ и $|y_n| \leqslant |z_n|$; и, с другой стороны, $|z_n| \leqslant |x_n| + |y_n|$.

В таком случае ясно следующее: *если сходится ряд* $\sum_{n=1}^{\infty} |z_n|$, *то сходится и ряд* $\sum_{n=1}^{\infty} z_n$. (Достаточно сослаться на теорему 1 на стр. 440 и на теорему a_1) на стр. 498.)

Таким образом, *исследование сходимости ряда с комплексными членами сводится к исследованию сходимости ряда с положительными членами.*

Абсолютно сходящиеся ряды обладают всеми свойствами конечных сумм: в частности, их сходимость не нарушается, и их сумма не изменяется, при каких угодно перестановках членов и при произвольной расстановке скобок; возможно также «умножение

ряда на ряд» по правилу «умножения многочлена на многочлен»[1]):

$$\sum_{n=1}^{\infty} z_n' \cdot \sum_{n=1}^{\infty} z_n'' = \sum_{p=1}^{\infty} \sum_{q=1}^{\infty} z_p' z_q''. \qquad (6)$$

Сказанное, вообще говоря, неверно по отношению к неабсолютно сходящимся рядам: в дальнейшем изложении, однако, мы будем избегать употребления таких рядов.

§ 3. Показательная функция. Синус и косинус

Школьный курс элементарной алгебры не создаёт никакого повода для того, чтобы ввести то или иное определение показательной функции при положительном основании и мнимых значениях показателя. Равным образом, тригонометрия не даёт никакого ответа на вопрос: что следует понимать под значением тригонометрических функций при мнимом значении аргумента?

Такой повод возникает, однако, после того как установлены формулы, содержащие разложения в степенные ряды

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots,$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots,$$

$$\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

Эти равенства справедливы при любых действительных значениях x (см. стр. 447) и, следовательно, являются тождествами в действительной области.

Правые части здесь обладают тем свойством, что они не теряют смысла также и при подстановке мнимых значений переменной: в самом деле, при такой подстановке получаются абсолютно сходящиеся ряды с комплексными членами, причём сумма каждого из них есть некоторое комплексное число[2]).

[1]) Доказательства этих свойств логически обязательны, так как законы действий не переносятся автоматически от случая конечного числа слагаемых на случай бесконечного числа (см. стр. 444—446).

[2]) Действительно, обозначая модуль подставляемого значения переменного через r, мы убеждаемся, что ряд

$$1 + \frac{r}{1!} + \frac{r^2}{2!} + \frac{r^3}{3!} + \dots$$

сходящийся (по признаку Даламбера; см. также стр. 450); и подавно сходящимися являются ряды

$$1 + \frac{r^2}{2!} + \frac{r^4}{4!} + \dots \quad \text{и} \quad \frac{r}{1!} + \frac{r^3}{3!} + \frac{r^5}{5!} + \dots,$$

получающиеся из него посредством выбрасывания части членов.

После этого уже не покажется удивительным, что в математике принято следующее о-п р е д е л е н и е:

Каково бы ни было значение (действительное или мнимое) *комплексного переменного z, под символами e^z, $\cos z$, $\sin z$ подразумевают суммы следующих абсолютно сходящихся рядов:*

$$\left.\begin{aligned} e^z &= 1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\dots,\\ \cos z &= 1-\frac{z^2}{2!}+\frac{z^4}{4!}-\frac{z^6}{6!}+\dots,\\ \sin z &= \frac{z}{1!}-\frac{z^3}{3!}+\frac{z^5}{5!}-\frac{z^7}{7!}+\dots \end{aligned}\right\} \qquad (7)$$

Или более кратко (и более исчерпывающе):

$$e^z=\sum_{n=0}^{\infty}\frac{z^n}{n!},\quad \cos z=\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n}}{(2n)!},\quad \sin z=\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n+1}}{(2n+1)!}. \qquad (8)$$

Таким образом, в с и л у о п р е д е л е н и я написанные здесь равенства являются тождествами во всей комплексной плоскости.

Необходимость именно этих (а не иных, им не равносильных) определений будет установлена — в известном смысле — позднее (см. ниже, § 15).

Каждая из формул (7 или 8) даёт возможность вычислять значение любой из трёх рассматриваемых функций приближённо, с наперёд заданной точностью; так, чтобы вычислить значение e^{3+2i} можно в правую часть первой из формул (7) подставить $3+2i$ вместо z и затем просуммировать достаточное число членов, разворачивая степени по формуле «бинома Ньютона».

Но вычисление можно произвести и гораздо проще, исходя из «теоремы сложения» — о с н о в н о г о с в о й с т в а п о к а з а т е л ь н о й ф у н к ц и и $f(z)=e^z$. Это свойство выражается равенством

$$f(z_1+z_2)=f(z_1)f(z_2), \qquad (9)$$

которое является тождеством относительно z_1 и z_2. В более простой записи

$$e^{z_1+z_2}=e^{z_1}e^{z_2} \qquad (10)$$

оно известно читателю для случая, когда z_1 и z_2 — действительные; но мы убедимся, что *оно справедливо и при произвольных комплексных значениях z_1 и z_2.*

Действительно, мы имеем:

$$f(z_1+z_2)=e^{z_1+z_2}=\sum_{n=0}^{\infty}\frac{(z_1+z_2)^n}{n!}=$$

$$=\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{p+q=n}\frac{n!}{p!q!}z_1^p z_2^q\ ^{1})=\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{1}{p!q!}z_1^p z_2^q=$$

$$=\sum_{p=0}^{\infty}\frac{z_1^p}{p!}\cdot\sum_{q=0}^{\infty}\frac{z_2^q}{q!}=\sum_{n=0}^{\infty}\frac{z_1^n}{n!}\cdot\sum_{n=0}^{\infty}\frac{z_2^n}{n!}=e^{z_1}e^{z_2}=f(z_1)f(z_2).$$

Из тождества (10) можно сейчас же вывести замечательное тождество Эйлера. Положим

$$z_1=x,\quad z_2=iy,\quad z_1+z_2=x+iy=z.$$

Тогда получится:

$$e^z=e^{x+iy}=e^x\cdot e^{iy} \tag{11}$$

Далее, по определению показательной функции

$$e^{iy}=1+\frac{iy}{1!}+\frac{(iy)^2}{2!}+\frac{(iy)^3}{3!}+\frac{(iy)^4}{4!}+\frac{(iy)^5}{5!}+\frac{(iy)^6}{6!}+\frac{(iy)^7}{7!}+\ldots=$$

$$=1+\frac{iy}{1!}-\frac{y^2}{2!}-\frac{iy^3}{3!}+\frac{y^4}{4!}+\frac{iy^5}{5!}-\frac{y^6}{6!}-\frac{iy^7}{7!}+\ldots=$$

$$=\left(1-\frac{y^2}{2!}+\frac{y^4}{4!}-\frac{y^6}{6!}+\ldots\right)+i\left(\frac{y}{1!}-\frac{y^3}{3!}+\frac{y^5}{5!}-\frac{y^7}{7!}+\ldots\right).$$

Но выражения в скобках по формулам (7) представляют собой не что иное, как $\cos y$ и $\sin y$. Итак,

$$e^{iy}=\cos y+i\sin y. \tag{12}$$

Это и есть тождество Эйлера: *оно выражает показательную функцию с чисто мнимым аргументом через косинус и синус* [2]) *с действительным аргументом.*

Мы вернёмся в следующем параграфе к тождеству Эйлера; но прежде подставим полученное выражение для e^{iy} в правую часть равенства (11) и получим формулу, по которой удобно вычислять

---

[1]) Здесь применена формула «бинома Ньютона» в следующей записи:

$$(z_1+z_2)^n=\sum_{p+q=n}\frac{n!}{p!q!}z_1^p z_2^q$$

(p и q пробегают целые неотрицательные значения, подчинённые условию $p+q=n$).

[2]) Мы предполагаем здесь и дальше, что буквами x и y обозначены действительные числа. Но ничто не препятствует рассматривать переменные x и y в формулах (11) и (12) так же, как произвольные комплексные числа.

значение показательной функции e^z при любом комплексном значении z:

$$e^z = e^x(\cos y + i \sin y). \tag{13}$$

Например, полагая $x = 3$, $y = 2$, мы будем иметь:

$$e^{3+2i} = e^3(\cos 2 + i \sin 2) \approx (2{,}71\ldots)^3(\cos 114°35' + i \sin 114°35') \approx$$
$$\approx -8{,}23 + 18{,}27i.$$

Формула (13), которая наравне с формулами (7) может служить определением показательной функции $w = e^z$ в комплексной области (и практически более удобна), обладает тем свойством, что позволяет сразу вычислить не только действительную часть u и мнимую часть v функции w, но также её модуль ρ и аргумент φ. Именно,

$$u = e^x \cos y, \quad v = e^x \sin y;$$

с другой стороны, так как выражение в скобках в правой части (13), очевидно, имеет модуль, равный единице, то легко понять, что

$$\rho = |e^z| = e^x, \quad \varphi = \arg e^z = y. \tag{14}$$

Из формулы Эйлера ясно видно замечательное (может быть, неожиданное для читателя) свойство показательной функции e^z: *функция e^z — периодическая*, она имеет *мнимый* период

$$\omega = 2\pi i.$$

Действительно, полагая в формуле (13) $y = 2\pi$, мы получаем:

$$e^{2\pi i} = 1, \tag{15}$$

и тогда тождество (10) при $z_1 = z$, $z_2 = \omega = 2\pi i$ даёт:

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z.$$

Таким же образом вообще при n целом:

$$e^{z+2n\pi i} = e^z. \tag{16}$$

Нам остаётся сказать несколько слов по поводу того случая, когда основанием показательной функции является произвольное положительное число a. Тогда по определению полагают:

$$a^z = e^{z \ln a}, \tag{17}$$

и отсюда с помощью формулы (13) следует:

$$a^z = a^x(\cos y \ln a + i \sin y \ln a). \tag{18}$$

Упражнения

1. Доказать, что $2^i \approx 0{,}77 + 0{,}64i$.

2. Дано $w = e^{z^2}$; доказать, что

$$u = e^{x^2-y^2} \cos 2xy, \quad v = e^{x^2-y^2} \sin 2xy.$$

3. Дано $w = ze^z$; доказать, что

$$u = e^x(x\cos y - y\sin y), \quad v = e^x(x\sin y + y\cos y).$$

4. Найти модуль и аргумент выражений

$$e^{3+2i},\ 2^i,\ 5^{3+2i},\ e^{\pi i},\ e^{\frac{\pi i}{2}}.$$

5. Дана функция $f(z) = \frac{\sin z}{z}$. Доказать, что $f(0) = 1$.

§ 4. Выражение тригонометрических функций через показательную

Вернёмся к формуле Эйлера, в которой заменим, впрочем, букву y буквой z; при этом будем предполагать, что z обозначает произвольное комплексное число

$$e^{iz} = \cos z + i\sin z. \tag{19}$$

Заменим, далее, z через $-z$:

$$e^{-iz} = \cos z - i\sin z.^1) \tag{20}$$

Решая полученные тождества[2]) (19) и (20) относительно величин $\cos z$ и $\sin z$, мы легко получим новые замечательные тождества:

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}. \tag{21}$$

Эти тождества выражают тригонометрические функции $\cos z$ *и* $\sin z$ *через показательную функцию.* Тот факт, что подобного рода выражение возможно, имеет громадное и принципиальное и практическое значение.

С принципиальной стороны, важно то, что в комплексной области тригонометрические функции теряют своё самостоятельное значение: можно всегда «без них обойтись», вводя вместо них показательные. Правда, это заключение — не для средней школы, где приходится довольствоваться обычным определением, относящимся только к действительной области, но зато обладающим свойствами наглядности и непосредственной применимости.

С практической стороны, существенно то, что, приняв в качестве определения тригонометрических функций формулы (21) и вводя вместо этих функций показательные, можно по большей части сократить записи и иногда также облегчить или автоматизировать вычисления. Приведём несколько примеров.

---

[1]) Свойства чётности косинуса и нечётности синуса

$$\cos(-z) = \cos z, \quad \sin(-z) = -\sin z$$

следуют из того, что степенное разложение косинуса (8) содержит лишь чётные степени, а разложение синуса (8) — лишь нечётные степени z.

[2]) См. подстрочное примечание [2]) на стр. 502.

1. Выведем «теорему сложения» для косинуса. Из тождества

$$\frac{e^{iz_1}+e^{-iz_1}}{2}\cdot\frac{e^{iz_2}+e^{-iz_2}}{2}-\frac{e^{iz_1}-e^{-iz_1}}{2i}\cdot\frac{e^{iz_2}-e^{-iz_2}}{2i}=\frac{e^{i(z_1+z_2)}+e^{-i(z_1+z_2)}}{2}$$

получаем сразу:

$$\cos z_1\cos z_2-\sin z_1\sin z_2=\cos(z_1+z_2). \tag{22}$$

(Аналогично для синуса.)

2. По теореме сложения показательной функции имеем:

$$e^{i(z_1+z_2+\ldots+z_n)}=e^{iz_1}e^{iz_2}\ldots e^{iz_n};$$

полагая все числа z_k ($k=1, 2, \ldots, n$) равными между собой и обозначая их общее значение через z, получим после перестановки правой и левой части:

$$(e^{iz})^n=e^{inz}. \tag{23}$$

Это — сокращённая запись известной формулы Муавра

$$(\cos z+i\sin z)^n=\cos nz+i\sin nz. \tag{24}$$

3. Пусть требуется представить $\cos^n z$ или $\sin^n z$ (где n — целое положительное) в виде линейной комбинации величин 1, $\cos z$, $\sin z$, $\cos 2z$, $\sin 2z$, ..., $\cos nz$, $\sin nz$.

Напишем, например, в случае косинуса:

$$\cos^n z=\left(\frac{e^{iz}+e^{-iz}}{2}\right)^n=\frac{1}{2^n}\{e^{inz}+C_n^1e^{i(n-2)z}+C_n^2e^{i(n-4)z}+\ldots+e^{-inz}\}=$$

$$=\frac{1}{2^{n-1}}\left\{\frac{e^{inz}+e^{-inz}}{2}+C_n^1\frac{e^{i(n-2)z}+e^{-i(n-2)z}}{2}+C_n^2\frac{e^{i(n-4)z}+e^{-i(n-4)z}}{2}+\ldots\right\}=$$

$$=\frac{1}{2^{n-1}}\{\cos nz+C_n^1\cos(n-2)z+C_n^2\cos(n-4)z+\ldots\}. \tag{25}$$

При этом последний член суммы в скобках равен

$$\begin{cases}\frac{1}{2}C_n^{\frac{n}{2}} & \text{в случае } n \text{ чётного,}\\ C_n^{\frac{n-1}{2}}\cos z & \text{в случае } n \text{ нечётного.}\end{cases}$$

4. Пусть требуется, напротив, выразить $\cos nz$ или $\sin nz$ через $\cos z$ и $\sin z$. Придётся написать, например, так:

$$\cos nz=\frac{e^{inz}+e^{-inz}}{2}=\frac{1}{2}\{(e^{iz})^n+(e^{-iz})^n\}=$$

$$=\frac{1}{2}\{(\cos z+i\sin z)^n+(\cos z-i\sin z)^n\},$$

и останется раскрыть скобки, пользуясь биномом Ньютона [1]).

[1]) Часто говорят несколько иначе: $\cos nz$ есть действительная часть от e^{inz}, т. е. от $(e^{iz})^n$, или $(\cos z+i\sin z)^n$; поэтому достаточно раскрыть скобки и «взять действительную часть» того, что получится. По существу разницы, конечно, нет.

5. Пусть нужно вычислить сумму

$$S_n = \frac{1}{2} + \cos z + \cos 2z + \ldots + \cos nz.$$

Вводя показательные функции, получаем геометрическую прогрессию, которую и суммируем по известной формуле:

$$S_n = \frac{1}{2} + \frac{e^{iz}+e^{-iz}}{2} + \frac{e^{2iz}+e^{-2iz}}{2} + \ldots + \frac{e^{niz}+e^{-niz}}{2} =$$
$$= \frac{1}{2}(e^{-inz} + \ldots + e^{-iz} + 1 + e^{iz} + \ldots + e^{inz}) =$$
$$= \frac{1}{2}\,\frac{e^{-inz} - e^{i(n+1)z}}{1 - e^{iz}} = \frac{1}{2}\,\frac{\sin\left(n+\frac{1}{2}\right)z}{\sin\frac{z}{2}}\ ^{1}).$$

Функция *тангенс* определяется как отношение синуса к косинусу; таким образом,

$$\operatorname{tg} z = \frac{\sin z}{\cos z} = \frac{1}{i}\,\frac{e^{iz}-e^{-iz}}{e^{iz}+e^{-iz}}\ . \tag{26}$$

Как известно, всякое комплексное число $z = x + iy$ $(\neq 0)$ можно представить «в тригонометрической форме»:

$$z = r(\cos\theta + i\sin\theta),$$

где $r = |z|$, $\theta = \arg z$; числа $r\,(>0)$ и θ определяются из уравнений

$$r\cos\theta = x,$$
$$r\sin\theta = y,$$

причём для r получается одно значение, а для θ — бесчисленное множество значений, отличающихся на величины, кратные 2π.

Основываясь на формуле Эйлера, условимся «тригонометрическую запись» ради краткости в дальнейшем заменять «показательной»:

$$z = re^{i\theta}. \tag{27}$$

Упражнения

1. Вычислить, чему равняется:

$$\cos i,\quad \cos(i\lg 2),\quad \sin\frac{1+i}{\sqrt{2}},\quad \operatorname{tg} i.$$

2. Решить уравнение

$$\sin z = 2.$$

3. Выделить действительную и мнимую части функций

$$\sin z,\quad \cos z,\quad \operatorname{tg} z.$$

---

[1]) Можно также вычислить сумму $\frac{1}{2} + e^{iz} + e^{2iz} + \ldots + e^{niz}$ и затем взять её действительную часть.

4. Проверить справедливость равенств, выражающих данные комплексные числа через их модуль и аргумент:

$$i = e^{i\frac{\pi}{2}}, \quad (-1) = e^{i\pi}, \quad 1+i = \sqrt{2}\, e^{i\frac{\pi}{4}},$$
$$1+i\sqrt{3} = 2e^{i\frac{\pi}{3}}, \quad 1-i\sqrt{3} = 2e^{-i\frac{\pi}{3}}.$$

§ 5. Гиперболические и тригонометрические функции

Известно (см. стр. 86), что функции

$$\left.\begin{aligned} \operatorname{ch} z &= \frac{e^z + e^{-z}}{2}, \\ \operatorname{sh} z &= \frac{e^z - e^{-z}}{2} \end{aligned}\right\} \tag{28}$$

носят названия *гиперболического косинуса* и *гиперболического синуса*; известны также аналогии, существующие между формальными свойствами тригонометрических и гиперболических функций.

Распространяя теперь определение (28) гиперболических функций на комплексную область, мы можем отдать себе отчёт в происхождении этих аналогий. Из сопоставления формул (21) и (28) следуют тождества

$$\left.\begin{aligned} \cos z &= \operatorname{ch}(iz), \\ i\sin z &= \operatorname{sh}(iz), \end{aligned}\right\} \tag{29}$$

или же [1])

$$\left.\begin{aligned} \operatorname{ch} z &= \cos(iz), \\ i\operatorname{sh} z &= \sin(iz). \end{aligned}\right\} \tag{30}$$

Таким образом, *гиперболические функции весьма просто выражаются через тригонометрические, и обратно*; отсюда и вытекают упомянутые аналогии.

Из всякого тождества, связывающего тригонометрические функции, можно вывести соответствующее тождество, связывающее гиперболические функции: достаточно представить себе, что вместо всякой буквы под знаком cos или sin подставлена та же буква с множителем i.

Например, заменяя в тождестве

$$\cos^2 z + \sin^2 z = 1$$

z через iz, получим тождество

$$\cos^2(iz) + \sin^2(iz) = 1, \quad \text{т. е.} \quad \operatorname{ch}^2 z - \operatorname{sh}^2 z = 1. \tag{31}$$

---

[1]) Формулы (30) можно получить также из формул (29) посредством замены z на $-iz$.

Аналогично из тождества

$$\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$$

следует тождество

$$\cos(iz_1 + iz_2) = \cos(iz_1)\cos(iz_2) - \sin(iz_1)\sin(iz_2),$$

т. е.

$$\operatorname{ch}(z_1 + z_2) = \operatorname{ch} z_1 \operatorname{ch} z_2 + \operatorname{sh} z_1 \operatorname{sh} z_2 \text{ [1]}. \tag{32}$$

Из формул (29) и (30), в частности, видно, что *при чисто мнимых значениях z функция* $\operatorname{ch} z$ *принимает те же значения, что и функция* $\cos z$ *при действительных значениях* z (и обратно); что касается функций $\operatorname{sh} z$ и $\sin z$, то имеется также различие в множителе i.

Гиперболические функции являются функциями периодическими, подобно тригонометрическим; но их период — не действительный (как у тригонометрических), а мнимый (как у функции e^z): он равен $2\pi i$.

Упражнения

1. Положив $z = x + iy$, вывести формулы

$$\cos z = \cos x \operatorname{ch} y - i \sin x \operatorname{sh} y,$$
$$\sin z = \sin x \operatorname{ch} y + i \cos x \operatorname{sh} y$$

и

$$\operatorname{ch} z = \operatorname{ch} x \cos y + i \operatorname{sh} x \sin y,$$
$$\operatorname{sh} z = \operatorname{sh} x \cos y + i \operatorname{ch} x \sin y.$$

2. Доказать, что корнями уравнения

$$\sin z = 0$$

в комплексной области являются числа вида $n\pi$, и т о л ь к о о н и.

3. Решить уравнение

$$\operatorname{ch} z = 0.$$

§ 6. Логарифм

Функция логарифм

$$w = \operatorname{Ln} z \text{ [2]}$$

в комплексной области, как и в действительной, может быть определена как о б р а т н а я по отношению к показательной $w = e^z$. Следовательно, для того чтобы узнать, чему равняется $w = \operatorname{Ln} z$ при данном значении z, надо решить относительно w уравнение

$$e^w = z. \tag{33}$$

Полагая $z = re^{i\theta}$, $w = u + iv$, мы придадим уравнению (33) вид

$$e^{u+iv} = re^{i\theta}. \tag{34}$$

---

[1]) Это тождество легко вывести и непосредственно из формул (28), не прибегая к мнимой единице (см. стр. 86).

[2]) В отличие от обыкновенного натурального логарифма положительного числа комплексный логарифм (сущность которого мы только выясняем) обозначается через $\operatorname{Ln} z$.

Если два комплексных числа равны, то модули их должны быть равны, а аргументы должны отличаться на величину, кратную 2π. Поэтому одно комплексное равенство (34) равносильно двум действительным:

$$\left.\begin{aligned} &e^u = r, \\ &v = \theta + 2k\pi \qquad (k \text{ — произвольное целое}). \end{aligned}\right\} \tag{35}$$

Отсюда следует (при $z \neq 0$), что $u = \ln r$ и, значит,

$$\operatorname{Ln} z = w = u + iv = \ln r + i(\theta + 2k\pi), \tag{36}$$

или

$$\operatorname{Ln} z = \ln |z| + i(\arg z + 2k\pi). \tag{37}$$

Таким образом, *логарифмом произвольного комплексного числа, не равного нулю, является всякое число, у которого* 1) *действительная часть равна обыкновенному логарифму от модуля данного числа,* 2) *мнимая часть равна одному из значений его аргумента.*

Отсюда следует прежде всего, что *всякое комплексное число* z $(\neq 0)$ *имеет бесчисленное множество логарифмов*: эти логарифмы образуют арифметическую прогрессию с разностью $2\pi i$. Если один из логарифмов числа z обозначим через w_0, то все логарифмы будут

$$\ldots, w_0 - 2n\pi i, \ldots, w_0 - 2\pi i, w_0, w_0 + 2\pi i, \ldots, w_0 + 2n\pi i, \ldots$$

В частности:

1) если z — действительное положительное число (аргумент θ равен нулю), то в качестве w_0 можно взять его обыкновенный (действительный) логарифм, и тогда все остальные логарифмы оказываются мнимыми;

2) если z — действительное отрицательное число или же мнимое число, то среди значений его аргумента θ нет кратных 2π, и потому все логарифмы без исключения — мнимые.

Заметим, наконец, что *число нуль не имеет ни одного логарифма,* так как уравнение $e^u = r$ при $r = 0$ не имеет корней.

Мы видим, таким образом, что введение комплексных чисел решительно изменяет (обобщает) понятие логарифма. Можно сказать, что во всякой области, не содержащей начала $z = 0$ [1]), функция $w = \operatorname{Ln} z$ бесконечно многозначна: меняя в формуле (37) целые значения k, мы переходим, как говорят, от одной «ветви» логарифма к другой.

Упражнения

1. Полагая $z = x + iy$, выразить $\operatorname{Ln} z$ через x и y.
2. Вычислить все логарифмы чисел:

$$1, -1, i, -i, 1+i, 1+i\sqrt{3}, 3+4i, 5-6i.$$

[1]) В этой точке функция $\operatorname{Ln} z$ «не существует».

§ 7. Произвольная степень

Произвольная степень $w = z^\alpha$, где α — какое угодно комплексное число, в комплексной области, как и в действительной, определяется через логарифм:

$$z^\alpha = (e^{\operatorname{Ln} z})^\alpha = e^{\alpha \operatorname{Ln} z} \qquad (z \neq 0). \tag{38}$$

Но разница, конечно, в том, что логарифм имеется в виду в обобщённом смысле, т. е. бесконечно многозначный. Таким образом, *функция* $w = z^\alpha$ также, вообще говоря, *обладает свойством бесконечной многозначности*. Именно, обозначая через $\ln z$ один из логарифмов числа z, мы получим:

$$\operatorname{Ln} z = \ln z + 2k\pi i,$$

и потому формула (38) примет вид

$$w = z^\alpha = e^{\alpha(\ln z + 2k\pi i)} = (e^{\ln z})^\alpha e^{2k\pi i\alpha} = w_0 e^{2k\pi i\alpha}.$$

Здесь w_0 — одно из значений функции w; все же эти значения образуют геометрическую прогрессию:

$$\ldots,\ w_0 e^{-2\pi i\alpha},\ \ldots,\ w_0 e^{-2\pi i\alpha},\ w_0,\ w_0 e^{2\pi i\alpha},\ \ldots,\ w_0 e^{2n\pi i\alpha},\ \ldots \tag{39}$$

Отметим, однако, частные случаи:

1) α — целое число: $\alpha = n$. Тогда $e^{2\pi i\alpha} = e^{2\pi in} = 1$, и все числа в строке (39) совпадают: *целая положительная степень комплексного числа* (как и следовало предвидеть) *имеет лишь одно значение.*

2) α — дробное рациональное число: $\alpha = \frac{p}{q}$ (несократимая дробь). Тогда среди чисел $e^{2k\pi i\alpha}$, т. е. среди чисел $e^{2k\pi i\frac{p}{q}}$, имеется ровно q различных: они соответствуют, например, значениям $k = 0, 1, 2, \ldots, q-1$; однако при $k = q$ мы уже получаем

$$e^{2k\pi i\frac{p}{q}} = e^{2\pi ip} = 1,$$

т. е. числа начинают повторяться. *Функция* $w = z^{\frac{p}{q}} = \sqrt[q]{z^p}$ [1]) *имеет,* таким образом, q *различных значений:* она — «q-значна». Её q значений, как легко убедиться, являются корнями уравнения $w^q = z^p$.

3) α — иррациональное действительное число. Тогда среди значений (39) нет равных между собой; все они в этом случае, как и в предыдущих, имеют один и тот же модуль.

4) α — чисто мнимое число: $\alpha = i\gamma$, где γ — действительное, $\gamma \neq 0$. Тогда знаменатель прогрессии (39) есть действительное положительное число:

$$e^{2\pi i\alpha} = e^{-2\pi\gamma} > 0;$$

---

[1]) Здесь мы вводим понятие корня целой положительной степени из комплексного числа.

все значения (39) различны между собой и, различаясь по модулю, имеют один и тот же аргумент.

5) $\alpha = \beta + i\gamma$, причём $\beta \neq 0$ и $\gamma \neq 0$. В этом наиболее общем случае все значения (39) различны между собой; притом могут меняться и модули и аргументы.

Упражнения

1. Доказать, что $i^i = e^{-\frac{\pi}{2}} e^{2k\pi}$ [1]).
2. Доказать, что $(-1)^{\sqrt{2}} = \cos(2k+1)\pi\sqrt{2} + i\sin(2k+1)\pi\sqrt{2}$.
3. Доказать, что $(-1)^i = e^{(2k+1)\pi}$.
4. Доказать, что

$$(1+i)^{1+i} = \sqrt{2}\, e^{-\left(\frac{\pi}{4}+2k\pi\right)}\left\{\cos\left(\ln\sqrt{2}+\frac{\pi}{4}\right)+i\sin\left(\ln\sqrt{2}+\frac{\pi}{4}\right)\right\}.$$

5. Полагая $z = x + iy = re^{i\theta}$, выделить действительную и мнимую части функции $w = z^i$; показать, что $|w| = e^{-(\theta+2k\pi)}$, $\arg w = \ln r$.
6. Сколько значений имеет выражение $(-1)^z$ при данном значении z: 1) целом, 2) дробном рациональном, 3) иррациональном действительном, 4) мнимом? Приведите примеры.

§ 8. Обратные тригонометрические и гиперболические функции

Тригонометрические и гиперболические функции весьма просто выражаются через показательную функцию (см. § 4); так как логарифм есть функция, обратная показательной, то неудивительно, что *функции, обратные тригонометрическим и гиперболическим, весьма просто выражаются через логарифм.*

Начнём с арксинуса. Понимая под символом

$$w = \operatorname{Arcsin} z$$

решение (или, лучше сказать, совокупность решений) уравнения

$$\sin w = z,$$

мы придадим этому уравнению вид

$$\frac{e^{iw} - e^{-iw}}{2i} = z,$$

что даёт нам сейчас же

$$e^{2iw} - 2ize^{iw} - 1 = 0.$$

---

[1]) Здесь и дальше k означает произвольное целое число.

Рассматривая последнее соотношение как квадратное уравнение относительно e^{iw}, мы получим [1]):

$$e^{iw} = iz + \sqrt{1 - z^2}.$$

Очевидно, iw есть один из логарифмов (любой) от выражения $iz + \sqrt{1 - z^2}$:

$$iw = \operatorname{Ln}(iz + \sqrt{1 - z^2});$$

отсюда, наконец, мы получаем:

$$w \equiv \operatorname{Arcsin} z = -i \operatorname{Ln}(iz + \sqrt{1 - z^2}). \tag{40}$$

Подобное же вычисление приводит к результату

$$\operatorname{Arccos} z = -i \operatorname{Ln}(z + i\sqrt{1 - z^2}). \tag{41}$$

И, наконец, из формулы (26) получаем:

$$\operatorname{Arctg} z = \frac{1}{2i} \operatorname{Ln} \frac{1 + iz}{1 - iz}. \tag{42}$$

Каждая из формул (40) — (42) бесконечно многозначна, так как бесконечно многозначен логарифм; кроме того, следует не упускать из виду, что радикалы в формулах (40) и (41) двузначны.

Чтобы получить функцию, обратную, например, косинусу гиперболическому, надо решить уравнение

$$\operatorname{ch} w = z,$$

которому, согласно (30), можно придать вид

$$\cos(iw) = z.$$

Отсюда

$$iw = \operatorname{Arccos} z,$$

и потому по формуле (41)

$$w = -i \operatorname{Arccos} z = -\operatorname{Ln}(z + i\sqrt{1 - z^2}),$$

или

$$\operatorname{Arch} z = \operatorname{Ln}(z + \sqrt{z^2 - 1}\,)\ ^{2}). \tag{43}$$

---

[1]) Уравнение $z^2 + pz + q = 0$ (где p и q — комплексные) решается в комплексной области, как и в действительной, посредством «выделения квадрата»: $z^2 + pz + \frac{p^2}{4} = \frac{p^2}{4} - q$, $\left(z + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q$, $z + \frac{p}{2} = \sqrt{\frac{p^2}{4} - q}$ и т. д. Знак радикала, ковечно, имеется в виду с двойным значением; операция извлечения корня у нас введена в § 7, п. 2).

[2]) В самом деле:

$$-\operatorname{Ln}(z + i\sqrt{1 - z^2}) = -\operatorname{Ln}(z + \sqrt{z^2 - 1}\,) = \operatorname{Ln} \frac{1}{z + \sqrt{z^2 - 1}} =$$

$$= \operatorname{Ln}(z - \sqrt{z^2 - 1}\,).$$

Однако безразличио, какой знак писать перед раднкалом, так как радикал — всё равно двузначный.

Подобным же образом

$$\operatorname{Arsh} z = -i \operatorname{Arcsin}(iz) = \operatorname{Ln}(z + \sqrt{z^2+1}), \tag{44}$$

$$\operatorname{Arth} z = -i \operatorname{Arctg}(iz) = \frac{1}{2} \operatorname{Ln} \frac{1+z}{1-z}. \tag{45}$$

Формулы (43) — (45) можно получить, конечно, и не прибегая к мнимой единице: достаточно искать обратные функции, исходя из выражений гиперболических функций через показательную (см. § 5).

Смысл выведенных здесь формул (43) — (45) всё же несколько иной, чем формул, выведенных на стр. 91—92, и именно потому, что 1) независимая переменная z предполагается имеющей произвольное комплексное значение, 2) логарифм имеется в виду «комплексный» (см. § 6), 3) радикалы в формулах (43) и (44) — двузначные; результатом последнего обстоятельства является то, что функция $\operatorname{Arsh} z$ при действительных значениях переменной ($z = x$), кроме значений вида $\ln(x + \sqrt{x^2+1}) + 2k\pi i$, имеет также значения вида

$$(2k+1)\pi i - \ln(x + \sqrt{x^2+1}).$$

Упражнения

1. Исходя из формул (43) — (45), напишите все значения каждого из выражений

$$\operatorname{Arcsin} 0, \quad \operatorname{Arccos} 0, \quad \operatorname{Arctg} 0, \quad \operatorname{Arcsin} 1, \quad \operatorname{Arccos} 1, \quad \lim_{z \to \infty} \operatorname{Arctg} z.$$

2. Считая z действительным ($z = x$), получаем в правой части формулы (42):

$$\frac{1}{2i} \operatorname{Ln} \frac{1+ix}{1-ix} = \frac{1}{2i} \left\{ \operatorname{Ln} \left| \frac{1+ix}{1-ix} \right| + i \arg \frac{1+ix}{1-ix} \right\} = \operatorname{Arctg} x.$$

Сделайте подобную же проверку по отношению к функциям $\operatorname{Arcsin} x$ и $\operatorname{Arccos} x$ (предполагая, конечно, что $|x| \leqslant 1$).

3. «Избавляясь» от тригонометрических и обратных тригонометрических функций (с помощью формул (21) и (41)), докажите, что многочлен Чебышева $T_n(z) = \cos n \operatorname{Arccos} z$ при всех значениях z может быть представлен в виде

$$T_n(z) = \frac{1}{2} \left[(z + \sqrt{z^2-1})^n + (z - \sqrt{z^2-1})^n \right].$$

§ 9. Производная

Говорят, что функция комплексного переменного $w = f(z)$, заданная и однозначная в некоторой окрестности данной точки z, имеет производную (дифференцируема), если существует и конечен предел

$$\lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z}. \tag{46}$$

Этот предел называется *производной* от функции $f(z)$ и обозначается $f'(z)$.

Вводя обозначения

$$\Delta z = h, \quad \Delta w = \Delta f(z) = f(z+h) - f(z),$$

можно также написать:

$$f'(z) = \lim_{\Delta z \to 0} \frac{\Delta f(z)}{\Delta z} = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \frac{dw}{dz}. \qquad (46')$$

Выражение $\frac{f(z+h)-f(z)}{h}$ есть функция от переменной h; здесь мы встречаемся с понятием предела функции. Формально в комплексной области это понятие определяется как в действительной; запись

$$\lim_{z \to z_0} F(z) = L,$$

или

$$F(z) \to L \quad \text{при} \quad z \to z_0,$$

означает, что, как бы мало ни было положительное число ε, можно указать такое δ, что из неравенства $|z - z_0| < \delta$ следует неравенство $|F(z) - L| < \varepsilon$; или (см. стр. 186) что, какова бы ни была последовательность $\{z_n\}$, подчинённая условию $z_n \to z_0$, непременно имеет место соотношение $F(z_n) \to L$.

Именно в этом смысле и следует понимать предельный переход в формуле (46).

Неравенство $|h| < \delta$ в действительной области означает, что точка $x+h$ лежит на отрезке оси Ox, образуемом точками $x-\delta$ и $x+\delta$. То же неравенство в комплексной области означает, что точка $z+h$ лежит внутри круга радиуса δ с центром z. Легко понять, что требование дифференцируемости несравненно стеснительнее в комплексной области, чем в действительной: предел (46) должен «не зависеть от того, каким образом точка $z+h$ приближается к точке z», причём в первом случае выбор путей, по которым $z+h$ приближается к z, гораздо разнообразнее, чем во втором.

В частности, допустим, что функция $f(z)$ задана в некоторой окрестности точки $z=x$ на действительной оси; тогда из дифференцируемости этой функции, понимаемой «в смысле теории функций действительного переменного», не вытекает дифференцируемость «в смысле теории функций комплексного переменного».

И тем не менее, *элементарные функции комплексного переменного* [1]) *оказываются дифференцируемыми всюду, за исключением разве лишь некоторых исключительных* («особых») *точек*, которые могут быть заранее выделены; притом *правила дифференцирования в комплексной области ничем не отличаются от правил дифференцирования в действительной области.*

---

[1]) Определение элементарных функций см. на стр. 11.

Последнее объясняется тем, что формальный аппарат, используемый при выводе правил дифференцирования, остаётся неизменным при переходе из действительной области в комплексную (хотя содержание формул меняется).

Рассмотрим простейший пример — правило дифференцирования функции $f(z)=z^2$. В комплексной области, как и в действительной, справедливо тождество (при $h\neq 0$)

$$\frac{f(z+h)-f(z)}{h}=\frac{(z+h)^2-z^2}{h}=2z+h$$

и далее, каким бы способом комплексное переменное приращение h ни приближалось к нулю, рассматриваемое выражение будет приближаться к пределу $2z$. Итак,

$$(z^2)'=2z.$$

Выводы правил дифференцирования основываются на теоремах о пределах; эти теоремы — на свойствах знака абсолютной величины (модуля); свойства же модуля в комплексной области — те же, что и свойства абсолютной величины в действительной (см. § 2).

Так сохраняются без изменений, вместе с их выводами, правила дифференцирования суммы, разности, произведения и дроби (частного).

Сохраняются также правило дифференцирования «функции от функции»

$$[f(\varphi(z))]'=f'(\varphi(z))\,\varphi'(z)$$

и правила дифференцирования «обратных» и «неявных» функций (мы не имеем возможности останавливаться на деталях, касающихся выделения «особых точек»).

Переходя к дифференцированию показательной функции и тригонометрических, следует не упустить из виду, что им было дано иное определение (§ 3), что влечёт за собой и необходимость пересмотра правил дифференцирования.

Из теоремы сложения для функции $f(z)=e^z$ следует:

$$\frac{f(z+h)-f(z)}{h}=\frac{e^{z+h}-e^z}{h}=e^z\cdot\frac{e^h-1}{h},$$

и остаётся ещё показать, что $\lim\limits_{h\to 0}\dfrac{e^h-1}{h}=1$, или что

$$\lim_{z\to 0}\frac{e^z-1}{z}=1. \tag{47}$$

Но это ясно из следующих соотношений:

$$\left|\frac{e^z-1}{z}-1\right|=\left|\frac{\left(1+\frac{z}{1!}+\frac{z^2}{2!}+\cdots\right)-1}{z}-1\right|=$$
$$=\left|\frac{z}{2!}+\frac{z^2}{3!}+\frac{z^3}{4!}+\cdots\right|\leqslant r\left(\frac{1}{2!}+\frac{r}{3!}+\frac{r^2}{4!}+\cdots\right),$$

так как, например, при $r<1$ выражение в последних скобках меньше, чем

$$\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\ldots=e-2,$$

и, следовательно, рассматриваемое выражение при $h\to 0$ стремится к нулю.

Итак,

$$(e^z)'=e^z. \tag{48}$$

Вывод правила дифференцирования синуса и косинуса основывается, как известно, на свойстве синуса

$$\lim_{z\to 0}\frac{\sin z}{z}=1; \tag{49}$$

нам нужно обобщить его на комплексную область. Мы получаем:

$$\left|\frac{\sin z}{z}-1\right|=\left|\frac{1}{z}\left(\frac{z}{1!}-\frac{z^3}{3!}+\frac{z^5}{5!}-\ldots\right)-1\right|\leqslant\frac{r^2}{3!}+\frac{r^4}{5!}+\ldots=$$
$$=r^2\left(\frac{1}{3!}+\frac{r^2}{5!}+\ldots\right);$$

дальнейшее ясно.

Впрочем, по свойствам степенных рядов возможно и непосредственное (почленное) дифференцирование [1]):

$$(e^z)'=\left(1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\ldots\right)'=\left(\frac{z}{1!}\right)'+\left(\frac{z^2}{2!}\right)'+\left(\frac{z^3}{3!}\right)'+\ldots=$$
$$=1+\frac{z}{1!}+\frac{z^2}{2!}+\ldots=e^z, \tag{50}$$

и аналогично

$$(\cos z)'=\left(1-\frac{z^2}{2!}+\frac{z^4}{4!}-\frac{z^6}{6!}+\ldots\right)'=$$
$$=-\frac{z}{1!}+\frac{z^3}{3!}-\frac{z^5}{5!}+\ldots=-\sin z, \tag{51}$$

$$(\sin z)'=\left(\frac{z}{1!}-\frac{z^3}{3!}+\frac{z^5}{5!}-\frac{z^7}{7!}+\ldots\right)'=1-\frac{z^2}{2!}+\frac{z^4}{4!}-\ldots=\cos z. \tag{52}$$

Функцию $w=\operatorname{Ln} z$ при выводе правила дифференцирования [2]) удобнее рассматривать как обратную по отношению к показательной:

$$e^w=z,\quad e^w w'=1,\quad w'=e^{-w},\quad \text{т. е.}\quad (\ln z)'=\frac{1}{z}. \tag{53}$$

Аналогично дифференцируются и функции, обратные по отношению к тригонометрическим.

---

[1]) Теоремы 2 и 3 (стр. 453—455) остаются в силе и для комплексной области (в основном сохраняются и доказательства).

[2]) Имеется в виду, конечно, какая-нибудь определённая (впрочем, безразлично какая) «ветвь» логарифма.

Мы не упомянули о свойстве *непрерывности* функций комплексного переменного. Это свойство определяется (как «непрерывность в точке», так и «непрерывность в области») буквально так же, как и в случае действительного переменного с заменой, конечно, «абсолютной величины» «модулем» (см. стр. 189).

Остаётся в силе, вместе с доказательством, теорема: *если функция дифференцируема* (в точке, в области), *то она непрерывна* (соответственно в точке, в области). См. стр. 309.

§ 10. Интеграл

Понятие определённого интеграла в комплексной области непосредственно обобщает соответствующее понятие в действительной области.

Пусть в плоскости $z = x + iy$ задана некоторая ориентированная кривая[1]) (L), идущая от начальной точки a до конечной точки b (рис. 1). Расставим на ней в порядке, соответствующем ориентации, точки z_k $(k = 1, 2, \ldots, n-1)$, причём положим также $z_0 \equiv a$, $z_n \equiv b$. Введём сокращённые обозначения

$$\Delta z_k = z_k - z_{k-1} \quad (k = 1, 2, \ldots, n).$$

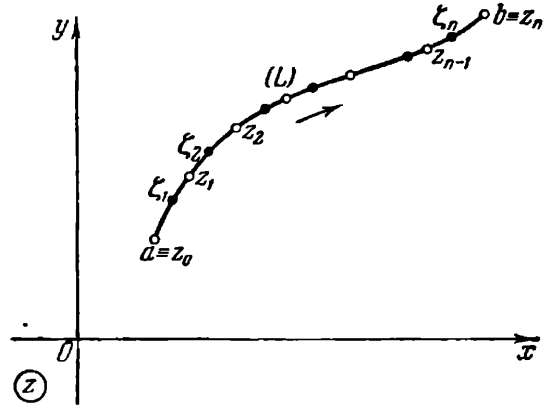


Рис. 1.

На каждой из частных дуг $z_{k-1}z_k$ возьмём произвольно ещё точку ζ_k (она может совпадать также и с каждым из концов (z_{k-1} или z_k)).

Предположим, что некоторая функция $w = f(z)$ задана во всех точках кривой (L); построим «интегральную сумму»:

$$S = \sum_{k=1}^{n} f(\zeta_k)(z_k - z_{k-1}) = \sum_{k=1}^{n} f(\zeta_k)\,\Delta z_k. \qquad (54)$$

Обозначим через δ наибольшее из положительных чисел $|\Delta z_k|$ — наибольшую из хорд ломаной линии с вершинами z_k.

Представим теперь, что от одного разбиения линий кривой (L) на части точками z_k $(k = 0, 1, \ldots, n)$ мы переходим к другому,

[1]) Говоря о «кривой», мы здесь имеем в виду такую кривую, которая либо в каждой точке имеет касательную, либо составлена из конечного числа дуг, имеющих касательную (таким образом, точки, где дуги соединяются между собой, оказываются угловыми). В частности, «кривой» может быть и ломаная линия.

от другого к третьему и т. д. — неограниченно. Может оказаться, что существует конечный предел I последовательности интегральных сумм, соответствующих рассматриваемым разбиениям, и что он не зависит от выбора последовательности разбиений — при единственном условии, что число δ будет стремиться к нулю (т. е. разбиения будут неограниченно мельчиться). Тогда предел I и называется *определённым интегралом от функции $f(z)$, взятым по дуге (L)*:

$$I = \lim_{\delta \to 0} S.$$

Обозначение определённого интеграла таково:

$$I = \int_{(L)} f(z)\, dz. \tag{55}$$

Таким образом, по определению мы имеем:

$$\int_{(L)} f(z)\, dz = \lim_{\delta \to 0} \sum_{k=1}^{n} f(\zeta_k)\, \Delta z_k. \tag{56}$$

Если функция $f(z)$ непрерывна на кривой (L), то конечный предел I существует (в указанном смысле), т. е. *функция $f(z)$ интегрируема по кривой (L).*

Доказательство приводит «комплексный» интеграл к действительному. Ограничимся для простоты предположением, что кривая (L) обладает тем свойством, что всякая прямая, параллельная действительной или мнимой оси, пересекает её в одной точке; предположим, другими словами, что уравнение кривой (L) может быть написано в любой из форм

$$y = y(x) \quad \text{или} \quad x = x(y),$$

причём $y(x)$ и $x(y)$ — непрерывные функции. Тогда, полагая

$$z_k = x_k + iy_k, \quad \Delta z_k = \Delta x_k + i\Delta y_k, \quad \zeta_k = \xi_k + i\eta_k,$$
$$f(z) \equiv u(x, y) + iv(x, y),$$

интегральную сумму S можно переписать в следующем виде:

$$S = \sum_{k=1}^{n} \{ u(\xi_k, \eta_k) + iv(\xi_k, \eta_k) \} (\Delta x_k + i\Delta y_k) =$$
$$= \sum_{k=1}^{n} \{ u(\xi_k, \eta_k)\, \Delta x_k - v(\xi_k, \eta_k)\, \Delta y_k \} +$$
$$+ i \sum_{k=1}^{n} \{ v(\xi_k, \eta_k)\, \Delta x_k + u(\xi_k, \eta_k)\, \Delta y_k \},$$

или иначе, принимая во внимание, что $\eta_k = y(\xi_k)$, $\xi_k = x(\eta_k)$,

$$S = \left\{\sum_{k=1}^{n} u(\xi_k, y(\xi_k))\,\Delta x_k - \sum_{k=1}^{n} v(x(\eta_k), \eta_k)\,\Delta y_k\right\} + i\left\{\sum_{k=1}^{n} v(\xi_k, y(\xi_k))\,\Delta x_k + \sum_{k=1}^{n} u(x(\eta_k), \eta_k)\,\Delta y_k\right\}.$$

Каждая из четырёх сумм в последней формуле есть интегральная сумма, составленная для непрерывных[1]) функций

$$u(x, y(x)), \quad v(x(y), y), \quad v(x, y(x)), \quad u(x(y), y) \tag{57}$$

по соответствующей переменной; каждая из этих сумм стремится поэтому при протекании интегрального процесса к пределу, равному соответствующему (действительному) интегралу. Отсюда ясно, что существует и (комплексный) интеграл $I = \lim S$.

Комплексные интегралы обладают, очевидно, следующими свойствами:

I. $$\int_{(L)} (af(z) + bg(z))\,dz = a\int_{(L)} f(z)\,dz + b\int_{(L)} g(z)\,dz \tag{58}$$

(a и b — произвольные, комплексные постоянные).

II. Если начальная точка кривой (L_2) совпадает с конечной точкой кривой (L_1) и через (L) обозначена кривая, составленная из кривых (L_1) и (L_2), то

$$\int_{(L)} f(z)\,dz = \int_{(L_1)} f(z)\,dz + \int_{(L_2)} f(z)\,dz. \tag{59}$$

III. Если кривая (L') от кривой (L) отличается только ориентацией (т. е. это та же кривая, но у неё b — начальная точка, a — конечная), то

$$\int_{(L')} f(z)\,dz = -\int_{(L)} f(z)\,dz. \tag{60}$$

Наконец,

IV. $$\left|\int_{(L)} f(z)\,dz\right| \leqslant LM, \tag{61}$$

где через M обозначено любое число, не меньшее чем $|f(z)|$, при каком бы то ни было значении z на дуге (L), а через L — длина дуги (L).

---

[1]) Легко понять, что если непрерывна функция $f(z)$ (на кривой (L)), то непрерывны также функции $u(x, y)$ и $v(x, y)$ (на этой кривой), а также и сложные функции (57) (см. стр. 190).

Это следует, в результате предельного перехода, из неравенства

$$|S| = \left|\sum_{k=1}^{n} f(\zeta_k)\,\Delta z_k\right| \leqslant \sum_{k=1}^{n} |f(\zeta_k)| \cdot |\Delta z_k| \leqslant M \sum_{k=1}^{n} |\Delta z_k|.$$

Последняя сумма, как легко понять, представляет собой периметр ломаной линии, вписанной в кривую (L).

Если роль дуги (L) играет отрезок действительной оси и если функция $f(z)$ принимает на нём лишь действительные значения, то комплексный интеграл обращается в действительный. Свойства I—IV являются обобщениями известных свойств действительного интеграла.

Мы условимся говорить, что *функция $f(z)$ интегрируема в некоторой односвязной области* (D) [1]), если эта функция интегрируема по любой кривой (L), принадлежащей области (D), и притом, каковы бы ни были две точки a и b из (D), имеет место равенство

$$\int_{(L_1)} f(z)\,dz = \int_{(L_2)} f(z)\,dz, \tag{62}$$

где (L_1) и (L_2) — произвольные кривые, принадлежащие (D) и идущие от a к b.

Легко понять, что последнее требование равносильно такому: какова бы ни была замкнутая [2]) кривая (C), принадлежащая (D), имеет место равенство

$$\int_{(C)} f(z)\,dz = 0. \tag{63}$$

Действительно, пусть интеграл по всякой замкнутой кривой равен нулю; в таком случае, если AmB и AnB — два пути, идущие от точки A к точке B (рис. 2), то

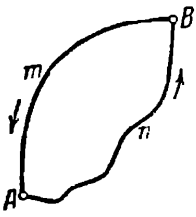


Рис. 2.

$$\int_{AnBmA} f(z)\,dz = 0,$$

т. е. (по свойству II)

$$\int_{BnA} f(z)\,dz + \int_{BmA} f(z)\,dz = 0$$

или (по свойству III)

$$\int_{AnB} f(z)\,dz - \int_{AmB} f(z)\,dz = 0, \quad \int_{AmB} f(z)\,dz = \int_{AnB} f(z)\,dz.$$

---

[1]) Необходимо уточнить, что́ следует в дальнейшем понимать под *односвязной областью* в комплексной плоскости. Это — или совокупность точек, расположенных внутри некоторой непрерывной замкнутой, себя не пересекающей, кривой, или совокупность точек, расположенных по одну сторону от некоторой кривой, идущей «из бесконечности в бесконечность» (вроде параболы), или вся плоскость. С дальнейшими обобщениями этого понятия нам не придётся иметь дело.

[2]) Это значит, что конечная точка кривой совпадает с начальной.

. Обратно, пусть значение интеграла по пути, идущему от одной данной точки к другой, не зависит от выбора пути. Пусть дан произвольный замкнутый контур; возьмём на нём две различные точки A и B и обозначим две образовавшиеся дуги контура через AnB и AmB. Тогда по предположению

$$\int_{AnB} f(z)\,dz = \int_{AmB} f(z)\,dz,$$

и отсюда следует (если перенесём второй интеграл налево и соединим интегралы):

$$\int_{AnBmA} f(z)\,dz = 0,$$

что и требовалось доказать

Всякая целая положительная степень $f(z) \equiv z^p$ интегрируема в любой области (D) (т. е. во всей плоскости).

В самом деле, раз кривая (C) — замкнутая, то в её разбиении конечная точка совпадает с начальной

$$z_n = z_0;$$

поэтому можно написать:

$$\sum_{k=1}^{n} z_k^{p+1} = \sum_{k=1}^{n} z_{k-1}^{p+1},$$

или

$$\sum_{k=1}^{n} (z_k^{p+1} - z_{k-1}^{p+1}) = 0,$$

или ещё

$$\sum_{k=1}^{n} (z_k^p + z_k^{\ 1} z_{k\ 1} + \quad + z_{k-1}^p)\,\Delta z_k = 0.$$

Этому соотношению можно также придать вид

$$\sum_{k=1}^{n} \left\{ \sum_{m=0}^{p} z_k^m z_{k-1}^{p-m} \right\} \Delta z_k = 0,$$

или

$$\sum_{m=0}^{p} \left\{ \sum_{k=1}^{n} z_k^m z_{k-1}^{p-m} \Delta z_k \right\} = 0.$$

Пусть теперь протекает обычный интегральный процесс; тогда каждая из сумм

$$S_m' \equiv \sum_{k=1}^{n} z_k^m z_{k-1}^{p-m} \Delta z_k \qquad (64).$$

стремится в пределе к интегралу $\int\limits_{(C)} z^p dz$ [1]), и так как число таких сумм равно $p+1$, то в пределе мы получаем

$$(p+1)\int\limits_{(C)} z^p dz = 0,$$

откуда сейчас же следует нужное нам равенство

$$\int\limits_{(C)} z^p dz = 0. \tag{65}$$

Ссылаясь теперь на свойство I интегралов, мы приходим к следующей важной теореме:

Интегрируемым в любой области (D), т. е. во всей плоскости, является также любой многочлен (целая рациональная функция)

$$f(z) = P(z)\ (\equiv az^n + bz^{n-1} + \ldots + kz + l).$$

Какие ещё функции являются интегрируемыми и в каких именно областях, мы увидим в следующем параграфе.

---

[1]) Сумма S'_m не является обычной интегральной суммой. Однако, сравнивая её с обычной интегральной суммой

$$S_m = \sum_{k=1}^{n} z_{k-1}^p \Delta z_k \qquad (\zeta_k = z_{k-1}),$$

мы получаем (обозначив через R максимум расстояния точек кривой (C) от начала):

$$|S'_m - S_m| = \left| \sum_{k=1}^{n} z_{k-1}^{p-m} (z_k^m - z_{k-1}^m) \Delta z_k \right| =$$

$$= \left| \sum_{k=1}^{n} z_{k-1}^{p-m} (z_k^{m-1} + z_k^{m-2} z_{k-1} + \ldots + z_{k-1}^{m-1}) \Delta z_k^2 \right| \leqslant$$

$$\leqslant \sum_{k=1}^{n} |z_{k-1}|^{p-m} (|z_k|^{m-1} + |z_k|^{m-2} |z_{k-1}| + \ldots + |z_{k-1}|^{m-1}) \cdot |\Delta z_k|^2 \leqslant$$

$$\leqslant mR^{p-1} \sum_{k=1}^{n} |\Delta z_k|^2.$$

Но

$$\sum_{k=1}^{n} |\Delta z_k|^2 = \sum_{k=1}^{n} |\Delta z_k| \cdot |\Delta z_k| < \sum_{k=1}^{n} \delta \cdot |\Delta z_k| = \delta \sum_{k=1}^{n} |\Delta z_k| \leqslant \delta C,$$

где C — длина кривой (C); итак, $|S'_m - S_m| < \delta C$.

Так как в интегральном процессе $\delta \to 0$, то разность $S'_m - S_m$ становится сколь угодно малой, откуда видно, что S'_m стремится к тому же пределу, что и S_m.

§ 11. Приближение функций многочленами

Пусть дана некоторая последовательность многочленов $\{P_n(z)\}$ ($n=1, 2, 3; \ldots$), равномерно сходящаяся на некоторой замкнутой кривой (C). Предел этой последовательности есть некоторая функция $f(z)$, заданная и непрерывная на кривой (C) [1]). Последовательность $\{P_n(z)\}$ и функция $f(z)$ связаны между собой таким образом, что, как бы мало ни было ε, можно указать такое N_ε, что при $n > N_\varepsilon$ имеет место неравенство

$$|P_n(z) - f(z)| < \varepsilon \tag{66}$$

для любой точки z на кривой (C).

В таком случае из неравенства (66) следует по свойству IV интеграла (см. § 10, IV: полагаем $M \equiv \varepsilon$):

$$\left|\int_{(C)} \{P_n(z) - f(z)\}\, dz\right| < \varepsilon C \quad (C — \text{длина кривой } (C)).$$

Так как интеграл $\int_{(C)} P_n(z)\, dz$ от многочлена по замкнутому контуру непременно равен нулю (см. конец § 10), то предыдущее неравенство принимает также вид

$$\left|\int_{(C)} f(z)\, dz\right| < \varepsilon C.$$

Здесь правая часть может быть сделана произвольно малой: значит, интеграл под знаком модуля равен нулю.

Итак, *при условии равномерной сходимости на замкнутой кривой* (C) *последовательности многочленов* $\{P_n(z)\}$ *к предельной функции* $f(z)$

$$P_n(z) \rightrightarrows f(z),$$

можно утверждать, что

$$\int_{(C)} f(z)\, dz = 0. \tag{67}$$

Допустим теперь, что последовательность многочленов $\{P_n(z)\}$ равномерно сходится к некоторой функции $f(z)$ в некоторой области (D); тогда равенство (67) имеет место для любой замкнутой кривой (C) в этой области, т. е. функция $f(z)$ интегрируема в области (D).

---

[1]) Понятие «функция» употребляется здесь (и дальше) в смысле соответствия: с каждой точкой z на кривой (C) сопоставляется некоторое число $f(z)$, в данном случае предел числовой последовательности $\{P_n(z)\}$. Что функция $f(z)$ непрерывна на (C), доказывается точно так же, как на стр. 519. Следует принять во внимание, что многочлены непрерывны в комплексной плоскости, так как они дифференцируемы в любой точке.

Приведём несколько примеров последовательностей многочленов указанного типа.

Пример 1.

$$P_n(z)=1+\frac{z}{1!}+\frac{z^2}{2!}+\ldots+\frac{z^n}{n!} \qquad (n=1,\ 2,\ \ldots).$$

Последовательность $\{P_n(z)\}$ сходится равномерно к функции $f(z)\equiv e^z$ в любой конечной области (D). Пусть все точки z области (D) находятся от начала координат на расстоянии, не большем чем R; тогда, принимая во внимание определение функции e^z (см. § 3), мы получаем в области (D):

$$|P_n(z)-e^z|=\left|\sum_{k=n+1}^{\infty}\frac{z^k}{k!}\right|\leqslant\sum_{k=n+1}^{\infty}\frac{R^k}{k!},$$

и выражение справа может быть сделано $<\varepsilon$ при n достаточно большом, так как оно является остаточным членом сходящегося числового ряда.

Пример 2.

$$\Pi_n(z)=\left(1+\frac{z}{n}\right)^n \qquad (n=1,\ 2,\ \ldots).$$

Сравним многочлены $\Pi_n(z)$ и $P_n(z)$; свободные члены и члены первой степени у них одинаковы, а коэффициент при z^k (если $k\geqslant 2$) у разности $P_n(z)-\Pi_n(z)$ равен

$$\frac{1}{k!}-\frac{C_n^k}{n^k}=\frac{1}{k!}\left[1-\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\right],$$

и это выражение меньше, чем

$$\frac{1}{k!}\left[1-\left(1-\frac{k-1}{n}\right)^{k-1}\right]<\frac{(k-1)^2}{k!n}\ ^{1})<\frac{1}{(k-2)!n};$$

поэтому при $|z|\leqslant R$

$$|\Pi_n(z)-P_n(z)|<\frac{1}{n}\sum_{k=2}^{n}\frac{R^k}{(k-2)!}<\frac{R^2}{n}\sum_{n=2}^{\infty}\frac{R^{k-2}}{(k-2)!}=\frac{1}{n}\cdot R^2e^R.$$

Если $n\to\infty$, то правая часть стремится к нулю; значит, во всякой ограниченной области $\Pi_n(z)$ стремится равномерно к тому же пределу, что и $P_n(z)$:

$$\left(1+\frac{z}{n}\right)^n\rightrightarrows e^z \quad (|z|\leqslant R)\ ^{2}).$$

---

[1]) В очевидном неравенстве (при $0<a<b$)

$$b^m-a^m<mb^{m-1}(b-a)$$

положим:

$$m=k-1,\quad a=1-\frac{k-1}{n},\quad b=1.$$

[2]) Сравните стр. 204.

Пример 3. Подобным же образом во всякой конечной области справедливы соотношения

$$\sum_{k=0}^{n}(-1)^k\frac{z^{2k}}{(2k)!}\rightrightarrows\cos z,\qquad\frac{\left(1+\frac{iz}{n}\right)^n+\left(1-\frac{iz}{n}\right)^n}{2}\rightrightarrows\cos z$$

(и аналогично для синуса).

Пример 4. Последовательность $\{z^n\}$ сходится к функции $f(z)\equiv 0$ в круге $|z|<1$, притом равномерно при условии $|z|\leqslant\rho$ (где $\rho<1$), так как

$$|z^n|\leqslant\rho^n\to 0.$$

Пример 5. Последовательность $\{1+z+z^2+\ldots+z^n\}$, т. е. ряд

$$1+z+z^2+\ldots+z^n+$$

сходится к функции $f(z)=\frac{1}{1-z}$ в том же круге, и равномерно при условии $|z|\leqslant\rho$ $(\rho<1)$, так как

$$\left|(1+z+\quad+z^n)-\frac{1}{1-z}\right|=\frac{|z|^{n+1}}{|1-z|}\leqslant\frac{\rho^{n+1}}{1-\rho}\to 0.$$

Пример 6. Последовательность

$$\left\{-\frac{A}{a-c}\left[1+\frac{z-c}{a-c}+\left(\frac{z-c}{a-c}\right)^2+\quad+\left(\frac{z-c}{a-c}\right)^n\right]\right\},$$

т. е. ряд

$$-\frac{A}{a-c}\left[1+\frac{z-c}{a-c}+\left(\frac{z-c}{a-c}\right)^2+\quad+\left(\frac{z-c}{a-c}\right)^n+\ldots\right]\ (c\neq a),$$

сходится к функции $f(z)=\frac{A}{z-a}$ внутри круга с центром c и радиусом, равным расстоянию между точками a и c, притом равномерно — во всяком концентрическом круге меньшего радиуса. Это вытекает из предыдущего результата, если только заменить z через $\frac{z-c}{a-c}$. При этом точка c может быть выбрана произвольно, $c\neq a$.

Мы будем говорить, что *функция $f(z)$ допускает приближение многочленами в некоторой точке z_0, если существует такая окрестность (D) точки z_0* (например, круг с центром в точке z_0), *что некоторая последовательность многочленов сходится к функции $f(z)$ равномерно в (D).*

Свойство функций допускать приближение многочленами имеет большое значение в теории функций комплексного переменного.

Если этим свойством обладают две данные функции (или любое конечное их число), то тем же свойством обладают также их сумма и их произведение: в сущности, это — теоремы о пределе «суммы» и «произведения» (см. стр. 169), с добавлением требования равномерности.

Из примера 6 вытекает, что «элементарная» дробная рациональная функция вида $\frac{A}{z-a}$ допускает приближение многочленами во всякой точке, кроме «полюса» a.

Но тогда уже можно заключить, что тем же свойством обладает любая рациональная функция; в самом деле, исключив из такой функции целую часть (многочлен), можно полученную дробную часть разложить на сумму «элементарных» дробей, из которых каждая обладает требуемым свойством [1]). Следует сделать лишь исключение для «полюсов» — тех точек, где знаменатель дроби обращается в нуль.

Из примеров 1—3 вытекает, что функции e^z, $\cos z$, $\sin z$ обладают интересующим нас свойством во всех точках плоскости.

Заслуживает внимания также теорема:

Если функция $\varphi(z)$ допускает приближение многочленами в точке z_0, а функция $f(w)$ в точке $w_0=\varphi(z_0)$, то функция $F(z)\equiv f(\varphi(z))$ допускает приближение многочленами в точке $z=z_0$.

Доказательство опирается, конечно, на тот факт, что «многочлен от многочлена» (сложная функция) есть также многочлен.

Так, например, функция $e^{\cos z}$ обладает интересующим нас свойством во всех точках плоскости; функция $e^{\frac{1}{z}}$ — во всех точках, кроме $z=0$.

§ 12. Первообразная функция

Основная теорема. *Если функция $f(z)$ допускает приближение многочленами в каждой точке некоторой односвязной области (D), то она в этой области интегрируема.*

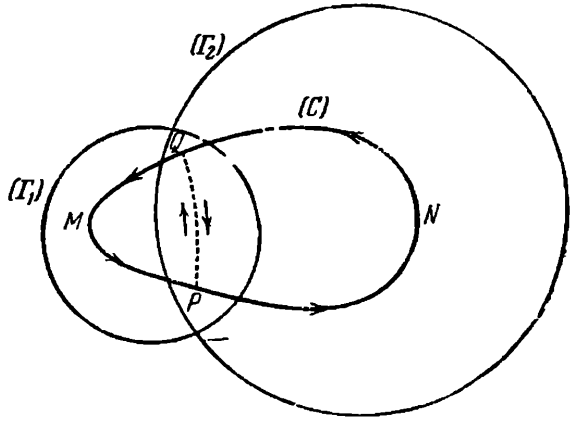


Рис. 3.

Мы уже рассмотрели (в начале § 11) тот случай, когда приближающая система многочленов одна и та же для всех точек рассматриваемой кривой интегрирования (L). Не останавливаясь на деталях общего доказательства, отметим, что оно основывается на следующем обстоятельстве.

[1]) В случае, если имеются кратные корни знаменателя дробн, необходимо провести дополнительное рассуждение: показать, что функции вида $\frac{A}{(z-a)^n}$ (n — целое, $\geqq 2$) допускают приближение многочленами во всякой точке, кроме $z=a$.

Если функция $f(z)$ интегрируема в каждом из двух налегающих друг на друга кругов (Γ_1) и (Γ_2), то она интегрируема и в области $(D)\equiv(\Gamma_1)+(\Gamma_2)$, составленной из точек, принадлежащих хоть одному из кругов (рис. 3). В самом деле, по предположению интеграл $\int f(z)\,dz$ равен нулю для всякой замкнутой кривой, заключённой внутри одного из кругов; рассмотрим теперь, например, кривую (C), изображённую на рис. 3 и умещающуюся в области (D), но не умещающуюся ни в одном из данных кругов. Имеем:

$$\int\limits_{(C)} f(z)\,dz=\int\limits_{(C_1)} f(z)\,dz+\int\limits_{(C_2)} f(z)\,dz,$$

где (C_1) — замкнутая кривая, составленная из дуги QMP кривой (C) и дополнительно проведённого пунктира PQ в «общей части» кругов (Γ_1) и (Γ_2), а (C_2) — такая же кривая, составленная из дуги PNQ кривой (C) и того же пунктира QP (ориентированного в другую сторону); так как каждый из интегралов справа равен нулю, то равняется нулю и интеграл слева.

Мы увидим из дальнейшего, что *условие односвязности области (D) является существенным.*

Если функция $f(z)$ задана, непрерывна и интегрируема в некоторой односвязной области (D), то интеграл $I=\int\limits_{(L)} f(z)\,dz$, взятый от точки a до точки b (в той же области), не зависит от выбора кривой интегрирования — лишь бы она не выходила за пределы области (D) (см. § 10). Поэтому при записи интеграла можно опускать название кривой и ставить вместо этого наименования начальной и конечной точек:

$$I=\int\limits_a^b f(z)\,dz.$$

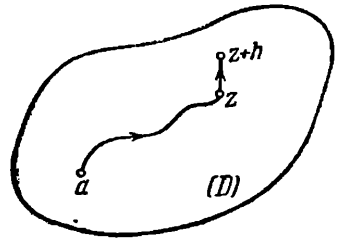


Рис. 4.

Мы убедимся сейчас, что *интеграл*

$$\Phi(z)=\int\limits_a^z f(\zeta)\,d\zeta\ ^{1}) \qquad (68)$$

(где a — произвольная постоянная точка из области (D), z — переменная точка из той же области, рис. 4) *представляет собой первообразную функцию от функции $f(z)$ в области (D)*, т. е. что имеет место тождество

$$\Phi'(z)=\frac{d}{dz}\int\limits_a^z f(\zeta)\,d\zeta=f(z)\ ^{2}). \qquad (69)$$

---

[1]) Интеграл не зависит от того, как обозначена «переменная интегрирования»; обозначение изменено, чтобы избежать смешения с верхним пределом.

[2]) Эта теорема обобщает теорему на стр. 391.

Действительно, мы имеем [1]):

$$\frac{\Phi(z+h)-\Phi(z)}{h}=\frac{1}{h}\left\{\int_a^{z+h} f(\zeta)\,d\zeta-\int_a^{z} f(\zeta)\,d\zeta\right\}=\frac{1}{h}\int_z^{z+h} f(\zeta)\,d\zeta=$$

$$=\frac{1}{h}\int_z^{z+h}\left[f(z)+(f(\zeta)-f(z))\right]d\zeta=$$

$$=\frac{1}{h}\left\{f(z)\int_z^{z+h} d\zeta+\int_z^{z+h}(f(\zeta)-f(z))\,d\zeta\right\}=$$

$$=f(z)+\frac{1}{h}\int_z^{z+h}(f(\zeta)-f(z))\,d\zeta.$$

Если точка $z+h$ достаточно близка к z, то второе слагаемое в последнем выражении меньше любого назначенного числа ε. В самом деле, выбирая прямолинейный путь интегрирования от z к $z+h$, применим свойство IV § 10 к интегралу

$$\int_z^{z+h}(f(\zeta)-f(z))\,d\zeta.$$

Здесь $L=|h|$, M равняется максимуму $|f(\zeta)-f(z)|$ при изменении ζ на отрезке от z до $z+h$; разность же $f(\zeta)-f(z)$ вследствие непрерывности $f(z)$ может быть сделана по модулю меньше любого числа ε, если только h достаточно мало. Итак,

$$\left|\frac{\Phi(z+h)-\Phi(z)}{h}-f(z)\right|<\varepsilon,$$

откуда и следует наше заключение.

В качестве примера (который поведёт к важным следствиям) рассмотрим интеграл

$$\Phi(z)=\int_1^z\frac{d\zeta}{\zeta}. \tag{70}$$

Функция $f(z)=\frac{1}{z}$ интегрируема во всякой односвязной области, не включающей начала; но мы не будем пока фиксировать области. Выберем такой путь интегрирования: сначала из точки 1 пойдём по действительной оси до точки $|z|=r$, затем от точки r — до точки z по дуге круга с центром в начале координат (рис. 5).

---

[1]) Предполагается, что точка $z+h$ находится в области (D) и что путь интегрирования от a к $z+h$ выбран проходящим через z (см. рис. 4).

Тогда

$$\Phi(z) = \int_1^r \frac{d\zeta}{\zeta} + \int_r^z \frac{d\zeta}{\zeta}.$$

В первом интеграле переменная пробегает положительные действительные значения $z = x$; поэтому

$$\int_1^r \frac{d\zeta}{\zeta} = \int_1^r \frac{dx}{x} = \ln r.$$

Полагая $z = re^{i\theta}$, $\zeta = re^{i\varphi}$ (см. рис. 5), возьмём во втором интеграле φ в качестве переменной интегрирования; тогда

$$d\zeta = ire^{i\varphi}\,d\varphi, \quad \frac{d\zeta}{\zeta} = i\,d\varphi,$$

и угол φ должен меняться от 0 до θ; итак,

$$\int_r^z \frac{d\zeta}{\zeta} = i\int_0^\theta d\varphi = i\theta.$$

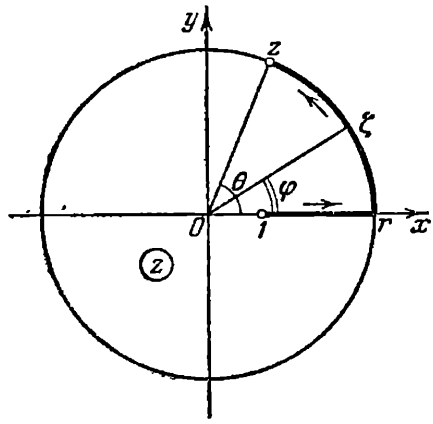


Рис. 5.

Замена переменной в этом комплексном интеграле нуждается в некоторых разъяснениях. Обратим внимание на то, что

$$e^z - 1 = z(1+\varepsilon), \qquad (71)$$

где величина ε, каково бы ни было z, при единственном ограничении $|z| \leqslant 1$ по модулю всегда меньше, чем z. Действительно,

$$\varepsilon = \frac{e^z - 1}{z} - 1 = \left(\frac{1}{2!} + \frac{z}{3!} + \frac{z^2}{4!} + \ldots\right) z,$$

и потому

$$|\varepsilon| < \left|\frac{1}{2!} + \frac{z}{3!} + \frac{z^2}{4!} + \ldots\right| \cdot |z| < \left(\frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots\right) \cdot |z| < (e-2)\cdot|z| < |z|.$$

Пусть на дуге круга между r и z расставлены точки

$$z_k = re^{i\varphi_k} \qquad (k = 0, 1, \ldots, n),$$

причём

$$0 < \varphi_1 < \ldots < \varphi_{n-1} < \varphi_n = \theta,$$

так что $z_0 = r$, $z_n = z$.

Тогда по определению интеграла (§ 10), полагая $\zeta_k = z_{k-1}$, получим:

$$\int_r^z \frac{d\zeta}{\zeta} = \lim \sum \frac{\Delta z_k}{z_{k-1}}.$$

Но, пользуясь равенством (71), мы видим, что

$$\Delta z_k = z_k - z_{k-1} = r\,(e^{i\varphi_k} - e^{i\varphi_{k-1}}) = re^{i\varphi_{k-1}}(e^{i\Delta\varphi_k} - 1) = ire^{i\varphi_{k-1}}\,\Delta\varphi_k\,(1+\varepsilon_k),$$

причём

$$|\varepsilon_k| < |\Delta\varphi_k| \qquad (k = 1, \ldots, n).$$

Далее,

$$z_{k-1} = re^{i\varphi_{k-1}},$$

и потому

$$\frac{\Delta z_k}{z_{k-1}} = i\Delta\varphi_k\,(1+\varepsilon_k).$$

Итак, интересующий нас интеграл $\int_r^z \frac{d\zeta}{\zeta}$ есть предел суммы

$$\sum_{k=1}^{n} i\,\Delta\varphi_k(1+\varepsilon_k) = i\sum_{k=1}^{n}\Delta\varphi_k + i\sum_{k=1}^{n}\varepsilon_k\,\Delta\varphi_k.$$

Первое слагаемое в правой части равно $i\theta$; что касается второго, то при условии, что $|\Delta\varphi_k| < \delta$, мы получаем $|\varepsilon_k| < |\Delta\varphi_k| < \delta$, и потому

$$\left| i\sum_{k=1}^{n}\varepsilon_k\,\Delta\varphi_k \right| \leqslant \sum_{k=1}^{n}|\varepsilon_k|\cdot|\Delta\varphi_k| < \delta\cdot\sum_{k=1}^{n}|\Delta\varphi_k| < \delta\cdot\theta;$$

таким образом, второе слагаемое стремится к нулю. Итак,

$$\int_r^z \frac{d\zeta}{\zeta} = i\theta.$$

Формулы, приведённые в основном тексте, следует рассматривать как краткую запись всего этого процесса.

Окончательно получаем:

$$\Phi(z) = \ln r + i\theta.$$

Этот результат как раз соответствует определению функции $\operatorname{Ln} z$ в комплексной области (см. § 6), однако здесь нет никакой неопределённости в аргументе. Если бы вместо того, чтобы из точки r сразу по дуге круга итти в точку z, мы предварительно сделали полный оборот (в прямом направлении) по кругу, то результат увеличился бы на выражение

$$i\int_0^{2\pi} d\varphi = 2\pi i;$$

если бы, вообще говоря, сделали k (целое число, $\gtrless 0$) таких оборотов, то получили бы дополнительное слагаемое $2k\pi i$. Таким путём можно получить все «ветви» комплексного логарифма (§ 6); многозначность его объясняется присутствием «полюса» $z = 0$, вокруг которого производится интеграция.

Если, в соответствии с изложенной выше теорией, ограничить изменение z какой угодно односвязной областью, не включающей этой точки, то многозначности, конечно, не возникает.

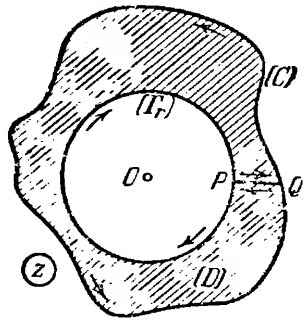


Рис. 6.

Мы убедились, что интеграл $\int \frac{d\zeta}{\zeta}$, взятый по кругу (Γ_r) произвольного радиуса r с центром в начале координат, равен $2\pi i$; вместе с тем взгляд на рис. 6 убеждает, что таково же значение этого интеграла, взятого по произвольной кривой (C) вокруг начала O. В самом деле, вообразим, что область (D), заштрихованная на рис. 6, «разрезана» по отрезку PQ, и станем интегрировать по контуру разрезанной области [1]): интеграл должен равняться нулю, так как «разрезанная» область (D) односвязна и не содержит начала; вместе с тем интегралы по отрезкам PQ и QP взаимно уничтожаются, и мы получаем

$$\int_{(C)} \frac{d\zeta}{\zeta} - \int_{(\Gamma_r)} \frac{d\zeta}{\zeta} = 0,$$

откуда следует:

$$\int_{(C)} \frac{d\zeta}{\zeta} = \int_{(\Gamma_r)} \frac{d\zeta}{\zeta} = 2\pi i. \tag{72}$$

Рассмотрим ещё интеграл $\int_{(C)} \frac{d\zeta}{\zeta - a}$, взятый по произвольной кривой, окружающей точку a — единственный полюс рациональной функции $f(z) = \frac{1}{z-a}$. В данном случае достаточно повторить прежнее рассуждение, вводя, однако, вместо кругов с центром в начале O, круги с центром в точке a и полагая $\zeta - a = re^{i\varphi}$. Значение интеграла окажется тем же самым, именно, $2\pi i$.

Заметим, наконец, что если точка a лежит в н е замкнутой кривой (C), то по основной теореме (см. начало § 12) интеграл $\int_{(C)} \frac{d\zeta}{\zeta - a}$ должен равняться н у л ю [2]).

Итак, мы приходим к важному заключению: *интеграл* $\int_{(C)} \frac{d\zeta}{\zeta - a}$, *взятый по замкнутой кривой* (C), *равен* $2\pi i$ или 0 *в зависимости*

---

[1]) Точнее говоря, здесь имеется в виду некоторый предельный переход.

[2]) В самом деле, функция $\frac{1}{z-a}$ допускает приближение многочленами во всякой точке, кроме точки $z = a$ (см. пример 6 в § 11).

от того, «охватывает» кривая (C) точку a или «не охватывает» [1]).

Упражнение.

Доказать (проводя рассуждения по образцу предыдущих), что интеграл $\int_{(C)} \frac{d\zeta}{(\zeta-a)^m}$, где m — произвольное целое число, отличное от единицы, равен нулю всегда — лишь бы кривая (C) не проходила через точку a.

§ 13. Интеграл Коши

Известно, что при делении произвольного многочлена $P(z)$ на $z-a$ (где a — произвольное комплексное число) в остатке получается число $P(a)$ [2]); таким образом, имеем тождество

$$P(z) \equiv (z-a)\,Q(z) + P(a), \tag{73}$$

в котором $Q(z)$ обозначает снова многочлен (частное при делении).

Частное от деления этого тождества на $z-a$ проинтегрируем по какой угодно замкнутой кривой, охватывающей точку a:

$$\int_{(C)} \frac{P(z)\,dz}{z-a} = \int_{(C)} Q(z)\,dz + P(a) \int_{(C)} \frac{dz}{z-a}.$$

Первый из интегралов в правой части равен нулю (§ 10); второй интеграл отличен от нуля (что чрезвычайно важно) и имеет значение $2\pi i$ (§ 12). Отсюда получаем, решая относительно $P(a)$:

$$P(a) = \frac{1}{2\pi i} \int_{(C)} \frac{P(z)\,dz}{z-a}. \tag{74}$$

Правая часть выведенной формулы (74), важнейшей в теории функций комплексного переменного, носит название *интеграла Коши*. Эта формула получена пока только для произвольного многочлена $P(z)$.

Однако формула (74) *имеет место и для любой функции $f(z)$, которая допускает приближение многочленами.* В самом деле, пусть последовательность многочленов $\{P_n(z)\}$ стремится равномерно к $f(z)$

$$P_n(z) \rightrightarrows f(z)$$

---

[1]) Интеграл «не имеет смысла», если кривая (C) проходит через точку a, так как функция $f(z) = \frac{1}{z-a}$ при этом «терпит разрыв» в точке a.

[2]) Так называемая «теорема Безу».

на кривой (C) и внутри неё. Переходя к пределу в формуле

$$P_n(a)=\frac{1}{2\pi i}\int\limits_{(C)}\frac{P_n(z)\,dz}{z-a},$$

мы получаем:

$$f(a)=\frac{1}{2\pi i}\int\limits_{(C)}\frac{f(z)\,dz}{z-a}\ ^{1}). \tag{75}$$

Посредством приёма, использованного в начале § 12 при доказательстве основной теоремы, дадим теперь результату дальнейшее обобщение — на функции $f(z)$, интегрируемые в некоторой области (D), при условии, что замкнутая кривая (C) находится в этой области и содержит внутри точку a (рис. 7).

Так как a — произвольная точка внутри (C), то её можно рассматривать как независимую переменную и обозначить буквой z; переменную же интегрирования обозначим буквой ζ. Тогда можно считать, что *интеграл Коши*

$$f(z)=\frac{1}{2\pi i}\int\limits_{(C)}\frac{f(\zeta)\,d\zeta}{\zeta-z} \tag{76}$$

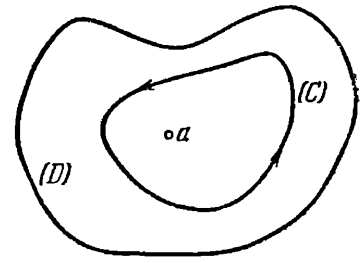


Рис. 7.

даёт представление значений функции $f(z)$ внутри (C) через её значения $f(\zeta)$ на самой кривой (C).

Посмотрим, какие следствия вытекают из факта представимости данной функции $f(z)$ интегралом Коши, взятым по данной кривой (C).

1. *Функция $f(z)$ дифференцируема внутри кривой* (C), причём для получения производной $f'(z)$ надо продифференцировать по z подинтегральную функцию:

$$f'(z)=\frac{1}{2\pi i}\int\limits_{(C)}\frac{f(\zeta)\,d\zeta}{(\zeta-z)^2}. \tag{77}$$

Для доказательства достаточно убедиться, что выражение

$$\frac{f(z+h)-f(z)}{h}$$

---

[1]) Переход к пределу в правой части обосновывается неравенствами

$$\left|\int\limits_{(C)}\frac{P_n(z)-f(z)}{z-a}\,dz\right|\leqslant C\max_{(C)}\left|\frac{P_n(z)-f(z)}{z-a}\right|\leqslant\frac{C}{\eta}\max_{(C)}\left|P_n(z)-f(z)\right|,$$

где η обозначает произвольное положительное число, не превышающее минимального расстояния точки a от точек кривой (C) (см. рис. 7).

стремится к правой части формулы (77) при $z_1 \to z$. Но это очевидно, так как

$$\frac{f(z+h)-f(z)}{h} = \frac{1}{2\pi i}\int_{(C)} \frac{f(\zeta)\,d\zeta}{(\zeta-z-h)(\zeta-z)}\ ^{1}).$$

2. *Функция $f(z)$ имеет внутри кривой (C) производную любого порядка n*, причём она получается посредством n-кратного дифференцирования по z подинтегральной функции:

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_{(C)} \frac{f(\zeta)\,d\zeta}{(\zeta-z)^{n+1}} \qquad (n=1,\ 2,\ \ldots). \tag{78}$$

Это легко доказывается по методу полной индукции: достаточно убедиться, что дифференцирование формулы (78) приводит к аналогичной формуле, в которой стоит $n+1$ вместо n [2]).

3. *Функция $f(z)$ разлагается в степенной ряд, расположенный по положительным степеням $z-c$, где c — произвольная точка внутри кривой (C). Этот ряд есть ряд Тейлора. Сходимость ряда к функции $f(z)$ имеет место по меньшей мере в круге с центром c и радиусом, не меньшим, чем расстояние точки c до ближайшей точки кривой (C).*

Всё дело в том, что в степенной ряд, расположенный по степеням $z-c$, разлагается функция $\frac{1}{\zeta-z}$, стоящая под знаком интеграла в формуле (76):

$$\frac{1}{\zeta-z} = \frac{1}{\zeta-c} + \frac{z-c}{(\zeta-c)^2} + \frac{(z-c)^2}{(\zeta-c)^3} + \ldots \tag{79}$$

(см. § 11, пример 6). В самом деле, мы имеем здесь прогрессию со знаменателем $\frac{z-c}{\zeta-c}$; если обозначим через η расстояние от точки c до ближайшей точки кривой (C) (рис. 8), то равномерная сходи-

---

[1]) Формально необходимо произвести «оценку» интеграла

$$\left|\frac{1}{2\pi i}\int_{(C)} \frac{f(\zeta)\,d\zeta}{(\zeta-z-h)(\zeta-z)} - \frac{1}{2\pi i}\int_{(C)} \frac{f(\zeta)\,d\zeta}{(\zeta-z)^2}\right| = \frac{|h|}{2\pi}\left|\int_{(C)} \frac{f(\zeta)\,d\zeta}{(\zeta-z)^2(\zeta-z-h)}\right|.$$

Если C обозначает длину кривой (C), M — наибольшее значение $|f(\zeta)|$ на (C), η — минимальное расстояние точки z от точек кривой (C), и если $|h| < \delta$, то рассматриваемое выражение меньше, чем

$$\frac{\delta}{2\pi}\cdot\frac{CM}{\eta^2(\eta-\delta)},$$

откуда легко вывести требуемое заключение.

[2]) При этом приходится произвести оценку интеграла, немногим более сложную, чем предыдущая; предоставляем её читателю.

мость разложения (79) обеспечивается условием $|z-c|<\theta\eta$, где θ — произвольное положительное число, меньшее единицы. Умножая равенство (79) на $f(\zeta)$, интегрируя [1]) по кривой (C) и деля на $2\pi i$, получим равенство:

$$\frac{1}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{\zeta-z}=\frac{1}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{\zeta-c}+$$
$$+\frac{z-c}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{(\zeta-c)^2}+\ldots$$
$$+\frac{(z-c)^n}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{(\zeta-c)^n}+\ldots \qquad (80)$$

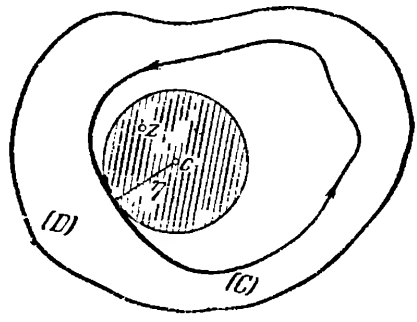


Рис. 8.

По формуле (76) выражение слева равняется $f(z)$; правой же части можно придать вид

$$a_0+a_1(z-c)+a_2(z-c)^2+\ldots\equiv\sum_{n=0}^{\infty}a_n(z-c)^n,$$

где положено

$$a_n=\frac{1}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{(\zeta-c)^{n+1}} \qquad (n=0,\ 1,\ 2,\ \ldots). \qquad (81)$$

Таким образом, равенство (80) можно записать в форме

$$f(z)=\sum_{n=0}^{\infty}a_n(z-c)^n. \qquad (82)$$

Мы получили разложение функции $f(z)$ в степенной ряд по степеням $z-c$. Равномерная сходимость обеспечивается условием

$$|z-c|<\theta\eta,$$

где $\theta\,(<1)$ как угодно мало отличается от единицы; иначе можно, следовательно, сказать; что возможность представления функции $f(z)$ рядом (82) обеспечивается условием $|z-c|<\eta$.

Сопоставляя, наконец, формулы (81) с формулами (76) и (78), мы убеждаемся, что

$$a_n=\frac{f^{(n)}(c)}{n!} \qquad (n=0,\ 1,\ 2,\ \ldots). \qquad (83)$$

Разложению (82) можно, таким образом, придать вид ряда Тейлора

$$f(z)=\sum_{n=0}^{\infty}\frac{f^{(n)}(c)}{n!}(z-c)^n. \qquad (84)$$

[1]) См. подстрочное примечание [1]) на стр. 516.

§ 14. Понятие аналитической функции

Исходя из понимания функции как формулы, содержащей в конечном числе лишь элементарные операции, мы приходим к сравнительно узкому классу элементарных функций комплексного переменного.

С другой стороны, самое общее понимание функции комплексного переменного

$$w = f(z) \qquad (z = x + iy,\ w = u + iv)$$

как соответствия между точками некоторой области в плоскости z и точками плоскости w приводит нас к рассмотрению соответствия между парами чисел (x, y) и парами чисел (u, v), т. е. системы двух действительных функций (в самом общем смысле) от двух действительных переменных

$$u = u(x, y),$$
$$v = v(x, y).$$

Это — объект изучения, достойный внимания, однако выходящий за пределы теории функций комплексного переменного.

Чтобы получить закономерное расширение понятия функции в комплексной области, поступают следующим образом: не покидая оперативной основы, добавляют к четырём рациональным операциям ещё только одну — операцию равномерного перехода к пределу. Возникающий при таком расширении класс функций носит название класса *аналитических функций*. Изучение свойств функций этого класса и составляет в собственном смысле предмет теории функций комплексного переменного. Свойства аналитических функций обладают замечательной взаимной обусловленностью, и совокупность (класс) этих функций в известном смысле «замкнут», т. е. не допускает дальнейшего расширения.

Высказанные мысли требуют уточнений и разъяснений.

Введём прежде всего понятие функции, *регулярной* в некоторой области. Что́ понимать под областью, об этом было сказано выше [1]); отметим всё же во избежание недоразумений, что под «областью» в теории функций комплексного переменного понимается множество точек плоскости, обладающее «двумерной протяжённостью»: например, отрезок прямой [2]) (или кривой) здесь никак нельзя назвать «областью».

---

[1]) См. подстрочное примечание на стр. 520. Научное определение области: *связное множество точек, для которого все его точки — «внутренние»*. Если ко множеству точек какой-нибудь области (например, внутренности круга) добавить хоть одну точку границы (в нашем примере — окружности), то вновь полученное множество перестаёт быть «областью».

[2]) То-есть как раз то, что называется «областью» или «промежутком» в теории функций действительной переменной.

Говорят, что функция $f(z)$, заданная (однозначно) и непрерывная в некоторой области (D), *регулярна* (или *правильна*, или *голоморфна*, или *аналитическая*) *в этой области*, если она обладает совокупностью свойств:

А) в каждой точке (D) функция $f(z)$ *допускает приближение многочленами*;

Б) в области (D) функция $f(z)$ *интегрируема*;

В) в каждой точке области (D) функция $f(z)$ *имеет производную*;

Г) в каждой точке c области (D) функция $f(z)$ *разлагается в ряд Тейлора*, равномерно сходящийся в некотором круге конечного радиуса с центром в точке c;

Смысл каждого из свойств А) — Г) был раскрыт раньше: А) — в § 11, Б) — в § 10, В) — в § 9 и Г) — в § 13.

Свойства А), Б), В) и Г) *равносильны между собой: каждое из них может быть логически выведено из каждого другого.*

В основу настоящего изложения было положено свойство А): мы видели, что из А) следует Б) (§ 12), также — с помощью интеграла Коши — следуют В) и Г) (§ 13).

Заметим, что из Г) немедленно следует А); в самом деле, раз функция разлагается в некотором круге в равномерно сходящийся степенной ряд, значит, частные суммы этого ряда, являющиеся многочленами, дают приближение функции.

С другой стороны, из В) следует Б). Мы не будем здесь приводить доказательство этого, более тонкого, предложения [1]). Далее, по этому ходу мыслей интеграл Коши получается [2]) следующим образом (без приближения многочленами). По предположению функция $f(z)$ имеет производную в рассматриваемой области; то же можно сказать о функции $\frac{1}{z-a}$ (кроме точки a), значит (§ 9), — и о произведении $\frac{f(z)}{z-a}$. В таком случае по свойству Б) произвольную кривую (C), охватывающую точку a, можно заменить окружностью (Γ_ρ) сколь угодно малого радиуса ρ с центром в этой точке; далее, получаем:

$$\frac{1}{2\pi i}\int\limits_{(C)}\frac{f(z)\,dz}{z-a}=\frac{1}{2\pi i}\int\limits_{(\Gamma_\rho)}\frac{f(z)\,dz}{z-a}=\frac{1}{2\pi i}\int\limits_{(\Gamma_\rho)}\frac{f(a)+[f(z)-f(a)]}{z-a}\,dz=$$

$$=f(a)+\frac{1}{2\pi i}\int\limits_{(\Gamma_\rho)}\frac{f(z)-f(a)}{z-a}\,dz.$$

Пользуясь непрерывностью $f(z)$, оцениваем последний интеграл и убеждаемся, что он по модулю меньше любого числа ε и, значит,

---

[1]) См. А. И. Маркушевич, Теория аналитических функций. Гостехиздат, 1950, гл. III, § 2.

[2]) Там же, § 3.

равен нулю. Отсюда получается интеграл Коши и, как его следствие, свойство Г).

Заслуживает особого внимания то обстоятельство, что из В) следует Г): *из того факта, что в некоторой области* (двумерной) *функция* $f(z)$ *имеет первую производную* $f'(z)$, *следует, что она «неограниченно дифференцируема»*, т. е. имеет производные $f^{(n)}(z)$ всех порядков n и, более того, *составленный из них ряд Тейлора*

$$f(c)+\frac{f'(c)}{1!}(z-c)+\frac{f''(c)}{2!}(z-c)^2+\ldots+\frac{f^{(n)}(c)}{n!}(z-c)^n+\ldots$$

сходится и имеет суммой $f(z)$ *в некоторой окрестности точки* c.

Ничего подобного нельзя утверждать относительно функций, заданных в действительной области, т. е. на отрезке.

Наконец, покажем, как из Б) непосредственно выводится В). Если в некоторой односвязной области интеграл $\int\limits_{(C)} f(\zeta)\,d\zeta$ не зависит от пути интегрирования, то наименование этого пути можно опустить (§ 10) и рассматривать интеграл $\int\limits_{z_0}^{z} f(\zeta)\,d\zeta$ как функцию z:

$$\Phi(z)=\int_{z_0}^{z} f(\zeta)\,d\zeta.$$

Но тогда (см. § 12) во всех точках рассматриваемой области функция $\Phi(z)$ имеет первую производную, именно,

$$\Phi'(z)=f(z),$$

а следовательно (см. предыдущий абзац), и вторую производную:

$$\Phi''(z)=f'(z).$$

Итак, функция $f(z)$ дифференцируема, что и следовало доказать [1]).

Чтобы установить регулярность функции в некоторой области, достаточно проверить, что выполнено о д н о из свойств А) — Г), тогда остальные следуют отсюда «автоматически». О регулярности элементарных функций легче всего судить по их дифференцируемости. Так, *элементарные функции регулярны во всех точках* [2]), *в окрестности которых они однозначно заданы и дифференцируемы* [3]).

---

[1]) Так называемая «теорема Морера».

[2]) Говорят: «функция регулярна в точке» вместо того, чтобы сказать «функция регулярна в некоторой окрестности точки» (например, в некотором круге с центром в этой точке).

[3]) Если функция задана неоднозначно, то выделяют её «однозначную ветвь» (как, например, в случае логарифма).

Примеры.

1. Функция $f(z)=\operatorname{tg} z$ регулярна во всей плоскости, кроме точек вида $\left(2n+1\right)\frac{\pi}{2}$.

2. Функция $f(z)=e^{\frac{1}{z}}$ регулярна во всей плоскости, кроме точки $z=0$.

Следующая теорема свидетельствует о замкнутости класса аналитических (регулярных в некоторой области) функций.

Всякая функция, регулярная в области, является равномерным пределом некоторой последовательности многочленов (свойство (A)). Регулярные функции, если можно так выразиться, «обобщают» многочлены. Возникает вопрос: рассматривая в некоторой области равномерно сходящуюся последовательность регулярных функций, не получим ли мы в пределе функцию, которая уже не будет регулярной (будет «обобщением» регулярной)? Ответ, оказывается, отрицательный: *если последовательность функций* $\{f_n(z)\}$, *регулярных в некоторой области* (D), *сходится в этой области равномерно, то функция, получающаяся в пределе,*

$$f(z)=\lim_{n\to\infty} f_n(z)$$

также регулярна в области (D) [1]).

Доказательство вытекает из интеграла Коши: взяв замкнутую кривую (C) в области (D), мы получаем при любом n

$$f_n(z)=\frac{1}{2\pi i}\int_{(C)}\frac{f_n(\zeta)\,d\zeta}{\zeta-z},$$

и переход к пределу даёт:

$$f(z)=\frac{1}{2\pi i}\int_{(C)}\frac{f(\zeta)\,d\zeta}{\zeta-z},$$

т. е. функция $f(z)$ представима (внутри (C)) интегралом Коши, а отсюда вытекает дифференцируемость и возможность разложения её в степенной ряд:

§ 15. Свойства аналитических функций

Особенно интересные свойства аналитических функций вытекают из их представимости степенным рядом (рядом Тейлора)

$$f(z)=\sum_{n=0}^{\infty} a_n(z-c)^n, \qquad a_n=\frac{f^{(n)}(c)}{n!}, \tag{85}$$

$$|z-c|<\rho, \quad \rho>0 \qquad (n=0,\ 1,\ 2,\ \ldots). \tag{86}$$

[1]) Теорема Вейерштрасса.

1. *Если функция $f(z)$ определена в круге сходимости* (86) *рядом* (85) *и, следовательно, регулярна в точке c и если сама она и все её производные в этой точке обращаются в нуль, то функция $f(z)$ в круге* (86) *тождественно равна нулю.*

Действительно, если

$$f^{(n)}(c)=0 \qquad (n=0,\ 1,\ 2,\ \ldots), \tag{87}$$

то тогда $a_n=0$ $(n=0,\ 1,\ 2,\ \ldots)$ и из формулы (85) следует:

$$f(z)\equiv 0.$$

В действительной же области, не вводя требования регулярности, можно построить функцию, удовлетворяющую условиям (87), и тождественно не равную нулю [1]).

2. *Для аналитических функций нули допускают классификацию по «кратностям».*

Говорят, что точка c есть нуль функции $f(z)$, если $f(c)=0$. Предположим, что функция $f(z)$ регулярна в точке c и что $f(c)=0$, но $f(z)\not\equiv 0$. Тогда существует такое целое положительное число p, что

$$f^{(n)}(c)=0 \text{ при } n=0,\ 1,\ \ldots,\ p-1, \text{ но } f^{(p)}(c)\neq 0:$$

---

[1]) Например, пусть $f(x)\equiv 0$ при $x\leqslant 0$, $f(x)\equiv e^{-\frac{1}{x^2}}$ при $x>0$.

Равенство $f^{(n)}(0)=0$ справедливо по определению при $n=0$, т. е. для самой функции $f(x)$. Убедиться в его справедливости при любом n можно посредством математической индукции. Пусть установлено, что $f^{(n)}(0)=0$; докажем, что $f^{(n+1)}(0)=0$. Мы имеем:

$$\frac{f^{(n)}(h)-f^{(n)}(0)}{h}=\frac{f^{(n)}(h)}{h}.$$

Это выражение, очевидно, равно нулю, если $h<0$; покажем, что при $h>0$

$$\lim_{h\to 0}\frac{f^{(n)}(h)}{h}=0.$$

С помощью той же индукции легко устанавливается, что $f^{(n)}(x)$ при $x>0$ имеет вид $e^{-\frac{1}{x^2}}P_n\left(\frac{1}{x}\right)$, где P_n — целый рациональный многочлен. Полагая и $P_n(u)\equiv Q_n(u)$, $\frac{1}{h}=u$, будем иметь при $h<1$, т. е. при $u>1$:

$$\frac{f^{(n)}(h)}{h}=e^{-u^2}Q_n(u)<e^{-u}Q_n(u),$$

откуда ясно (вследствие свойства показательной функции, доказанного на стр. 81—82), что правая часть, а значит, и левая, стремятся к нулю вместе с h. Итак,

$$\lim_{h\to 0}\frac{f^{(n)}(h)-f^{(n)}(0)}{h}=0,$$

как бы h ни стремилось к нулю, пробегая значения, отличные от нуля; а это как раз и означает, что $f^{(n+1)}(0)=0$.

иначе по предыдущему мы получили бы $f(z)\equiv 0$. Число p называется *кратностью* нуля c функции $f(z)$. Если c есть нуль кратности p, то формула (85) принимает вид

$$f(z)=(z-c)^p\left\{a_p+\sum_{n-p+1}^{\infty}a_n(z-c)^{n-p}\right\}\quad\left(a_p=\frac{f^{(p)}(c)}{p!}\neq 0\right),$$

и потому имеет место тождество

$$f(z)\equiv(z-c)^p f_1(z), \tag{88}$$

где $f_1(z)$ — функция, регулярная в точке c [1]), и не обращается в ней в нуль; справедливо, как легко понять, и обратное утверждение: если функция $f(z)$ имеет вид (88), то она имеет нуль p-й кратности в точке c.

Понятие кратности нуля общеизвестно в случае многочленов; но оно распространяется, таким образом, и на все аналитические функции.

3. *Нули аналитических функций* ($\not\equiv 0$) *изолированы.* Это значит: если функция $f(z)$, тождественно не равная нулю, регулярна в точке c и в ней имеет нуль, то можно указать круг $|z-c|<\delta$ ($\delta>0$), в котором, кроме c, нет больше нулей $f(z)$.

Пусть p — кратность нуля c. Утверждение вытекает из формулы (88). Функция $f_1(z)$ непрерывна (так как регулярна) в точке c и в ней отлична от нуля (так как $f_1(c)=a_p\neq 0$); выберем δ настолько малым, чтобы при $|z-c|<\delta$ имело место неравенство $|f_1(z)-f_1(c)|<\varepsilon\equiv|f_1(c)|$; последнее же влечёт за собой $f_1(z)\neq 0$ [2]) (при $|z-c|<\delta$).

4. *Если две функции $f(z)$ и $g(z)$, обе регулярные в некоторой области (D), принимают одинаковые значения во всех точках некоторой области (D_1), представляющей собой часть области (D), или же во всех точках некоторой кривой (C), целиком принадлежащей (D), то во всей области (D) имеет место тождество* $f(z)\equiv g(z)$.

В противном случае функция

$$h(z)\equiv f(z)-g(z),$$

тождественно не равная нулю в (D), имела бы нулями все точки частной области (D_1) или кривой (C), и тогда её нули не были бы изолированными.

Последнее из указанных свойств позволяет обосновать, в частности, выбор определения функций e^z, $\cos z$, $\sin z$ для комплексной области. Так, мы приняли (§ 3) в качестве определения, что под

---

[1]) Потому что представляется степенным рядом.

[2]) Из $|b-a|<|a|$, $a\neq 0$ следует $b\neq 0$.

значением показательной функции e^z в комплексной области (D) понимается сумма ряда

$$f(z) \equiv 1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots \qquad (89)$$

Эта сумма:

1) представляет собой регулярную функцию z во всей плоскости,

2) при действительных значениях переменной $z = x + iy$ (т. е. при $y = 0$) обращается (см. стр. 471) в известную из элементарной математики показательную функцию действительного переменного x

$$f(x) \equiv e^x \qquad (-\infty < x < +\infty).$$

Из свойства 4 вытекает, что *функция $f(z)$, заданная рядом* (89), — *единственная, удовлетворяющая требованиям* 1) — 2) или даже менее стеснительным требованиям — быть регулярной в неко-

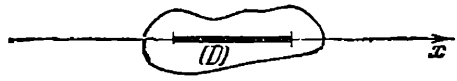


Рис. 9.

торой области (D), заключающей внутри себя какой-нибудь отрезок действительной оси (рис. 9), и обращаться на этом отрезке в функцию e^x.

Приведём ещё другой пример применения свойства 4.
В тригонометрии доказывается тождество

$$\sin(a+b) = \sin a \cos b + \cos a \sin b \qquad (90)$$

при условиях $a > 0$, $b > 0$, $a + b < \frac{\pi}{2}$ и затем посредством ряда последовательно проводимых рассуждений «распространяется» на всевозможные действительные значения a и b. Однако мы можем теперь «распространить» его сразу на все комплексные значения входящих букв. Пусть сначала b — действительное, $0 < b < \frac{\pi}{2}$. Рассмотрим функцию

$$f(z) \equiv \sin(z+b) - \{\sin z \cos b + \cos z \sin b\}. \qquad (91)$$

Она, как легко понять, регулярна во всей плоскости и притом вследствие (90) тождественно равна нулю на отрезке $0 < x < \frac{\pi}{2} - b$, значит, она тождественно равна нулю во всей плоскости; отсюда следует справедливость равенства (90) при $0 < b < \frac{\pi}{2}$ и всевозможных комплексных значениях a. В таком случае функция $f(z)$, будучи симметричной относительно z и b, обращается в нуль при произвольном комплексном значении b и при действительных значениях z, удовлетворяющих неравенству $0 < z < \frac{\pi}{2}$; но тогда она обращается в нуль и при всех комплексных значениях z; последнее означает, что равенство (90) выполняется тождественно во всей плоскости.

Вот ещё более тонкий пример. Считая доказанным равенство

$$a^{\frac{1}{p}} \cdot a^{\frac{1}{q}} = a^{\frac{1}{p}+\frac{1}{q}}$$

при всех целых положительных значениях p и q, мы, немедленно «распространяя» его, получаем тождество (для комплексных значений z_1 и z_2)

$$a^{z_1} a^{z_2} = a^{z_1+z_2},$$

опираясь на то, что функция

$$f(z) \equiv a^z \, a^{\frac{1}{q}} - a^{z+\frac{1}{q}},$$

очевидно, регулярная во всей плоскости, имеет нули вида $\frac{1}{p}$ с предельной точкой $z=0$.

Свойство 4 лежит в основе *аналитического продолжения* функций.

Предположим, что так или иначе заданы в некоторой области (D) значения функции $f(z)$; эти значения таковы, что $f(z)$ оказывается регулярной в (D). Предположим, что (D) не совпадает со всей плоскостью и что (D_1) — некоторая область, более «обширная», чем (D), т. е. содержащая её, но не совпадающая с нею (рис. 10). Можно поставить вопрос: существует ли функция $f_1(z)$, обладающая свойствами:

1) $f_1(z)$ регулярна в (D_1),

2) $f_1(z) \equiv f(z)$ в (D)?

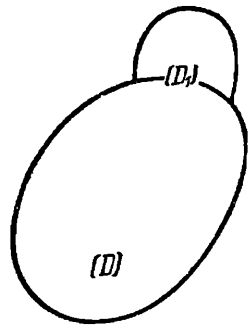


Рис. 10.

Такой функции может не существовать; но если она существует, то она единственна (свойство 4). Это даёт основание отождествлять функцию $f_1(z)$ с функцией $f(z)$, считая значения $f_1(z)$, взятые в точках области (D_1), но вне области (D), значениями $f(z)$ в этих точках.

Принимая такого рода «принцип аналитического продолжения», теория функций комплексного переменного разрешает говорить о значениях функции в некоторых точках даже в том случае, если эти точки лежат за пределами области определения функции.

С другой стороны, принято говорить о значениях функции комплексного переменного лишь в тех точках, в которых функция регулярна (см. § 14).

Таким образом, в теории функций комплексного переменного считается, что аналитическая функция, помимо той области, в которой она непосредственно задана тем или иным способом, определяется также во всех областях, в которые она может быть аналитически продолжена: из всех этих областей, вместе взятых, и составляется *область существования* аналитической функции.

Возможность «ветвления» функции и возникновения многозначности не противоречит принципу аналитического продолжения. Так, может случиться, что из области $(D)\equiv abcda$ с двойной штриховкой (рис. 11) функция продолжена аналитически в область $(D_1)\equiv aefghibcda$, заштрихованную вертикально, и в область $(D_2)\equiv abcjhkfda$, заштрихованную горизонтально, причём значения функции в области $fghkf$ при одном и при другом продолжении окажутся различными.

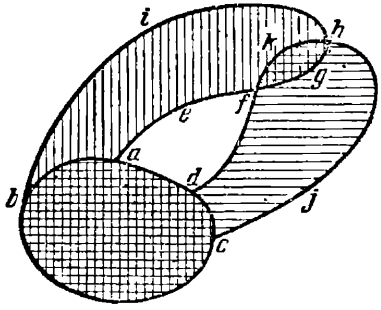


Рис. 11.

Следующие примеры иллюстрируют резкое различие взглядов на природу функции с точки зрения теории функций действительного переменного (ТФДП) и с точки зрения теории функций комплексного переменного (ТФКП).

Пример 1. Задана функция действительного переменного $y=x^2$ в промежутке $-1\leqslant x\leqslant+1$. Согласно ТФДП она существует в этом промежутке и нигде более. Согласно ТФКП она «продолжается» не только на всю действительную ось, но и на всю комплексную плоскость с помощью формулы $w=z^2$, так как это — единственно возможное аналитическое продолжение.

Пример 2. Задана функция действительного переменного $y=|x|$ в промежутке $-1\leqslant x\leqslant+1$. С точки зрения ТФДП здесь имеется о д н а функция, заданная в названном промежутке. С точки зрения ТФКП аналитическое продолжение из правой половины отрезка на всю плоскость даёт функцию $w=z$, из левой — функцию $w=-z$: здесь имеется, таким образом, д в е различные аналитические функции.

Пример 3. Задано уравнение $y^2=x$. Согласно ТФДП оно определяет д в е различные (непрерывные) функции на полуоси $0\leqslant x<\infty$: $y=\sqrt{x}$ и $y=-\sqrt{x}$ (радикалы — арифметические). Согласно ТФКП уравнение определяет о д н у аналитическую функцию $w=\sqrt{z}$ (радикал — алгебраический), имеющую по два значения для всех комплексных значений $z(\neq 0)$: одна «ветвь» переходит в другую в результате аналитического продолжения «вокруг» особой точки $z=0$.

Пример 4. То же можно сказать о функции $y=\sqrt{1-x^2}$ — решении уравнения $x^2+y^2=1$ (см. стр. 15 и 219). С точки зрения ТФКП мы получаем только одну аналитическую функцию $w=\sqrt{1-z^2}$ (радикал—алгебраический), имеющую по два значения для всех комплексных значений, кроме $z=1$ и $z=-1$, причём одна «ветвь» переходит в другую при обходе каждой из точек $+1$ и -1.

§ 16. Геометрический смысл аналитических функций

Как мы видели (см. § 1), задавая в некоторой области (D) совершенно произвольную функцию комплексного переменного $w=f(z)$, мы тем самым сопоставляем с каждой точкой z области (D) некоторую точку в плоскости w; можно сказать, что функция $w=f(z)$ «порождает» некоторое отображение области (D) в плоскости z на плоскость w.

Нас интересует теперь вопрос: какое геометрическое свойство порождаемого отображения соответствует свойству регулярности функции $f(z)$ в области (D)?

Свойство регулярности $f(z)$ в области (D) может быть характеризовано (§ 15) тем, что в каждой точке z этой области функция $f(z)$ имеет производную $f'(z)$. Посмотрим, что означает геометрически этот последний факт: выясним отдельно, каков геометрический смысл а) модуля производной $|f'(z)|$, б) аргумента производной $\arg f'(z)$.

Положим, как обычно,

$$\Delta z = h, \quad \Delta w = \Delta f(z) = f(z+h) - f(z)$$

и, кроме того,

$$\Delta z = \rho e^{i\varphi},$$

так что

$$|\Delta z| = \rho, \quad \arg \Delta z = \varphi.$$

Допустим, что точка $z+h$ приближается к точке z по некоторому л у ч у, т. е. таким образом, что аргумент φ приращения Δz

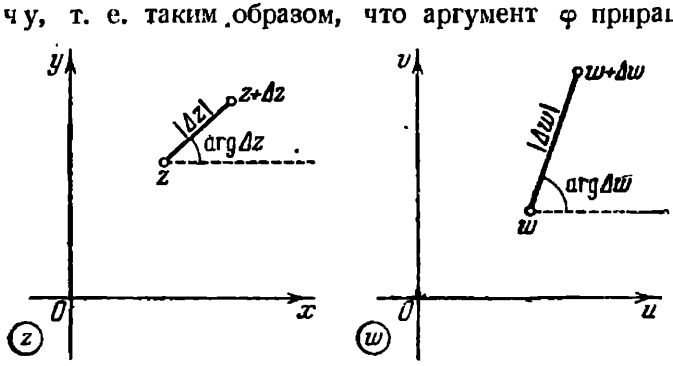


Рис. 12.

остаётся постоянным, а модуль его ρ стремится к нулю (рис. 12). Допустим, что в точке z производная w' отлична от нуля.

По определению производной предел

$$w' = \lim \frac{\Delta w}{\Delta z}$$

существует, конечен и не зависит от того, каким образом Δz стремится к нулю. Это означает, что оба следующих предела

$$\text{а)}\ |w'| = \lim \frac{|\Delta w|}{|\Delta z|} \quad \text{и} \quad \text{б)}\ \arg w' = \lim \{ \arg \Delta w - \arg \Delta z \}$$

существуют, конечны и не зависят, в частности, от угла φ [1]).

Отношение $\frac{|\Delta w|}{|\Delta z|}$ показывает, во сколько раз расстояние между точками $w+\Delta w$ и w больше, чем расстояние между точками

[1]) См. мелкий шрифт в § 2.

$z+\Delta z$ и z; назовём его предел $|w'|$ *масштабом* рассматриваемого отображения (в точке z).

Разность $\arg \Delta w - \arg \Delta z$ показывает, на какой угол нужно повернуть вектор, имеющий начало в точке z и конец в точке $z+\Delta z$, чтобы направление его совпало с направлением вектора, имеющего начало в точке w и конец в точке $w+\Delta w$ (см. рис. 12); назовём предел этого угла $\arg w'$ *кручением* рассматриваемого отображения (в точке z).

Итак, *если функция* $w=f(z)$ *регулярна в области* (D), *то*, какова бы ни была точка z из (D), *масштаб и кручение* в этой точке отображения, порождаемого функцией $f(z)$, *не зависят от того, по какому направлению точка* $z+\Delta z$ *приближается к точке* z.

Представим себе теперь, что по разным направлениям к точке z приближаются одновременно (независимо одна от другой) две точки: $z_1 \equiv z+\Delta z_1$ и $z_2 \equiv z+\Delta z_2$. Отметим в плоскости w соответствующие точки w, $w_1 \equiv w+\Delta w_1$ и $w_2 \equiv w+\Delta w_2$ (рис. 13). Точки w_1 и w_2 также будут приближаться к точке w, но уже не обязательно по прямым лучам, а может быть, по кривым, имеющим в точке w некоторые касательные (wT_1 и wT_2).

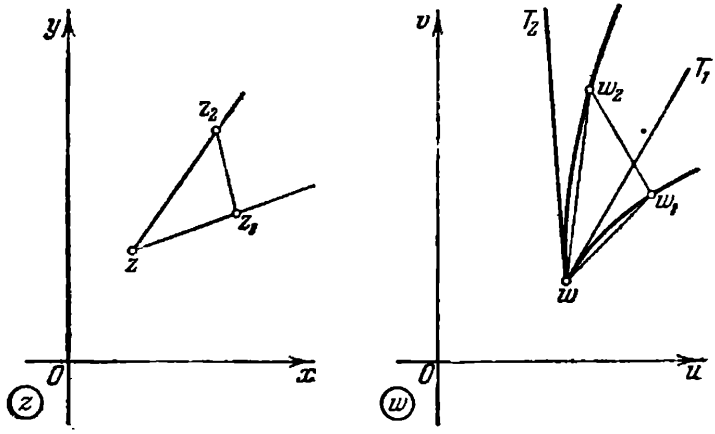


Рис. 13.

Рассмотрим треугольники: zz_1z_2 в одной плоскости и ww_1w_2 — в другой.

По свойству постоянства масштаба, отношения сторон ww_1 к zz_1 и ww_2 к zz_2 в пределе будут равны, т. е. между сторонами обнаружится пропорциональность.

По свойству постоянства кручения будут равны между собой углы между лучами zz_1 и wT_1, zz_2 и wT_2, т. е. будут равны углы z_1zz_2 и T_1wT_2.

Таким образом, «*в пределе*» *треугольники* zz_1z_2 и ww_1w_2 *становятся подобными*; при отображении, порождаемом функцией

$w=f(z)$, как говорят, *имеет место подобие в «бесконечно малой» окрестности каждой точки области* (D).

Это замечательное свойство отображений, порождаемых регулярными функциями, носит название *конформности*.

Можно доказать, что, обратно, *всякое конформное отображение некоторой области* (D) *в плоскости* z *порождается*[1] *некоторой функцией* $w=f(z)$, *регулярной в этой области*.

Таким образом, *конформность порождаемого отображения есть ещё одно из свойств, способных характеризовать аналитические функции* (см. § 14).

§ 17. Примеры конформных отображений

Пример 1. $w=z+c$.

Отображение представляет собой параллельное перенесение на вектор Oc.

В самом деле, положив $c=a+ib$ и разделяя действительную и мнимую части в данном уравнении, мы получаем:

$$\begin{cases} u=x+a, \\ v=y+b, \end{cases}$$

что как раз соответствует формулам параллельного переноса в аналитической геометрии.

Пример 2. $w=mz$, где $|m|=1$, $m=e^{i\omega}$.

Отображение — поворот около начала O на угол ω.

Действительно, разделение действительной и мнимой частей в данном уравнении

$$w=e^{i\omega}z$$

даёт нам теперь известные из аналитической геометрии формулы поворота:

$$\begin{cases} u=x\cos\omega-y\sin\omega, \\ v=x\sin\omega+y\cos\omega. \end{cases}$$

Пример 3. $w=mz$, где m — положительное число.

Отображение — растяжение в m раз около начала по направлению обеих осей (т. е. «увеличение» в m раз, «подобие без вращения», «гомотетия»), так как мы получаем:

$$\begin{cases} u=mx, \\ v=my. \end{cases}$$

---

[1]) При дополнительном условии сохранения «ориентировки» плоскости (т. е. сохранения направления при обходе по замкнутой кривой).

Пример 4. $w = mz$, где m — произвольное, отличное от нуля комплексное число.

Если положим

$$m = \mu e^{i\omega} \qquad (\mu > 0),$$

то, вводя «промежуточное» комплексное переменное t, будем иметь два последовательно выполняемых отображения:

$$w = \mu t,$$
$$t = e^{i\omega} z.$$

Таким образом, точка t получается из точки z посредством поворота на угол ω, а точка w — из точки t посредством растяжения в μ раз.

Возникающее «сложное» отображение иногда называют «локсодромическим» («подобие с вращением»).

Пример 5. $w = mz + n$, причём $|m| = 1$, $m = e^{i\omega}$.

Это отображение также «разлагается» на два:

$$w = t + n,$$
$$t = e^{i\omega} z.$$

При переходе от точки z к точке t совершается поворот на угол ω, затем при переходе от точки t к точке w — параллельный перенос на вектор On.

Очевидно, данное преобразование представляет собой общий случай движения на плоскости.

Пример 6. $w = mz + n$, где $m \neq 0$.

Полагая $m = \mu e^{i\omega}$, «разложим» отображение на три:

$$w = t_2 + n,$$
$$t_2 = \mu t_1,$$
$$t_1 = e^{i\omega} z.$$

В итоге получается комбинация из поворота, растяжения и параллельного переноса, т. е. отображение представляет собой подобие (общий случай).

Пример 7. $w = \frac{1}{z}$.

Разделение действительной и мнимой частей приводит к формулам:

$$\begin{cases} u = \dfrac{x}{x^2 + y^2}, \\ v = \dfrac{-y}{x^2 + y^2}. \end{cases}$$

Сопоставляя их с формулами, имеющимися на стр. 262, мы видим, что рассматриваемое отображение — инверсия относительно

круга $|z|=1$, однако с последующим симметричным отражением относительно оси Ox.

Полагая $z=re^{i\theta}$, $w=Re^{i\varphi}$, можно также написать формулы:

$$\begin{cases} R=\frac{1}{r}, \\ \varphi=-\theta, \end{cases}$$

яснее передающие геометрическое содержание отображения.

Пример 8. $w=z^{\alpha}$ $(\alpha>0)$.

Вводя также и на этот раз модули и аргументы обеих переменных, будем иметь:

$$\begin{cases} R=r^{\alpha}, \\ \varphi=\alpha\theta. \end{cases}$$

Отображение, очевидно, таково, что окружность с центром в начале на плоскости z превращается в окружность же с центром

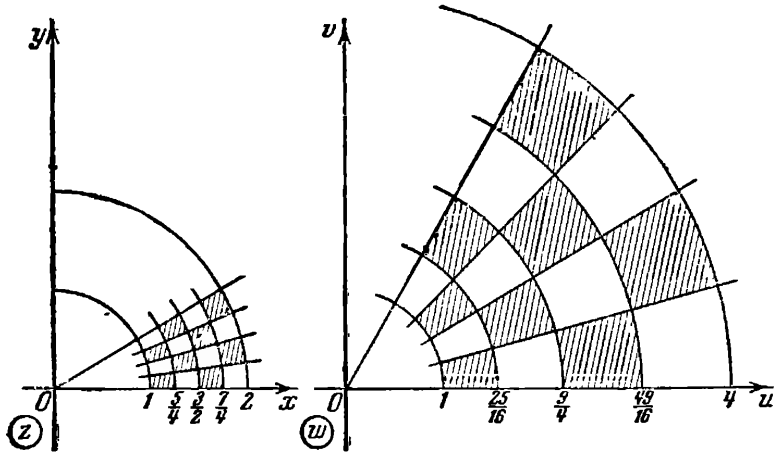


Рис. 14.

в начале на плоскости w, причём изменение радиусов происходит в зависимости от первого уравнения; второе же уравнение говорит о том, что лучи, выходящие из начала на плоскости z, переходят в такого же рода лучи на плоскости w, но угол луча с начальным лучом (т. е. аргумент произвольной точки на луче) всякий раз увеличивается в α раз.

На рис. 14 показана часть полярной сетки на плоскости z и её отображение на плоскость w (при $\alpha=2$).

Пример 9. $w=\operatorname{Ln} z$.

Положим $z=re^{i\theta}$, $w=u+iv$ (это равносильно введению полярных координат на плоскости z и декартовых — на плоскости w).

Тогда получим (см. § 6):

$$\begin{cases} u = \ln r, \\ v = \theta + 2k\pi \quad (k \text{ — произвольное целое}). \end{cases}$$

Наличие в этих уравнениях произвольной целочисленной постоянной k свидетельствует о бесконечной многозначности отображения плоскости z на плоскость w.

Приняв, например, $k = 0$, мы получаем отображение всей плоскости z с «выбросом» точки $z = 0$, т. е. области

$$0 < r < \infty, \quad 0 \leqslant \theta < 2\pi,$$

на «полосу» в плоскости w, определяемую неравенствами

$$-\infty < u < +\infty, \quad 0 \leqslant v < 2\pi;$$

именно, окружности вида $r = r_0$ переходят в прямые $u = \ln r_0$, а лучи $\theta = \theta_0$ — в прямые $v = \theta_0$.

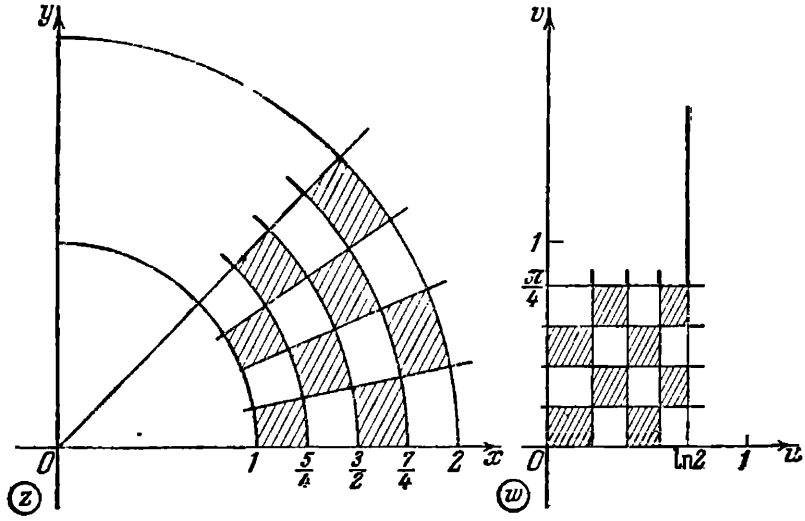


Рис. 15.

Положив, далее, $k = 1$, получим аналогичное отображение той же плоскости на «параллельную» полосу в плоскости w

$$-\infty < u < +\infty, \quad 2\pi \leqslant v < 4\pi,$$

и т. д. — для всех целых k.

Лéгко понять, что одна и та же точка $z = z_0$ ($\neq 0$) в плоскости z отображается на двустороннюю арифметическую прогрессию точек в плоскости w с разностью $2\pi i$.

На рис. 15 изображена часть полярной сетки в плоскости z и её отображение (в виде прямоугольной сетки) на плоскость w.

Пример 10. $w = \sin z$.

В данном случае, наоборот, удобнее ввести прямоугольные координаты на плоскости z и полярные — на плоскости w:

$$z = x + iy, \qquad w = Re^{i\varphi}.$$

Мы получаем после подстановки:

$$Re^{i\varphi} = \sin(x + iy),$$

или (см. § 5)

$$Re^{i\varphi} = \sin x \operatorname{ch} y + i \cos x \operatorname{sh} y.$$

Отсюда следует:

$$\begin{cases} R = \sqrt{(\sin x \operatorname{ch} y)^2 + (\cos x \operatorname{sh} y)^2}, \\ \varphi = \operatorname{Arctg} \dfrac{\cos x \operatorname{sh} y}{\sin x \operatorname{ch} y}, \end{cases}$$

т. е.

$$\begin{cases} R = \sqrt{\operatorname{ch}^2 y - \cos^2 x}, \\ \varphi = \operatorname{Arctg} \dfrac{\operatorname{th} y}{\operatorname{tg} x}. \end{cases}$$

Последние уравнения дают возможность вычислить модуль и аргумент $w = \sin z$ для каждого данного значения z.

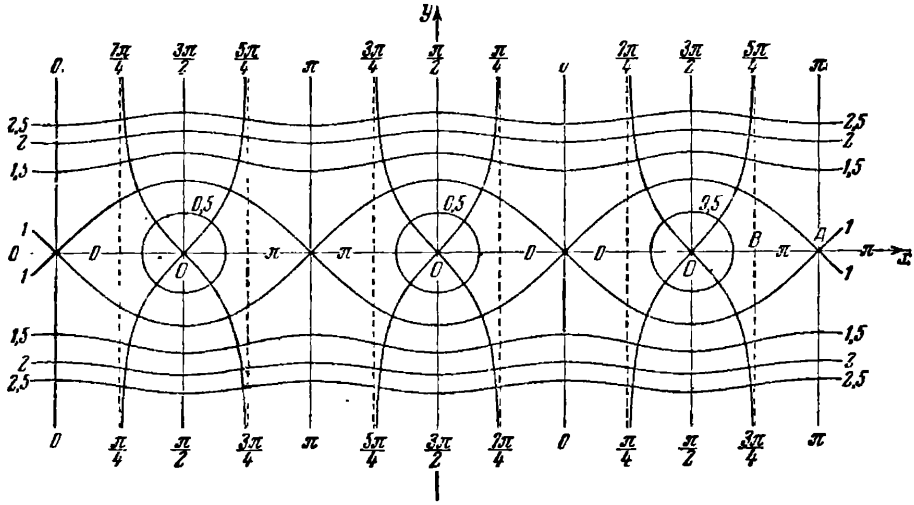


Рис. 16.

Решая эти уравнения относительно y, найдём геометрическое место точек в плоскости z, для которых: а) $|\sin z|$ имеет данное значение R ($\neq 0$) или б) $\arg \sin z$ имеет данное значение φ ($\neq 0$):

$$\text{а)}\ y = \pm \ln\left(\sqrt{R^2 + \cos^2 x} + \sqrt{R^2 - \sin^2 x}\right),$$

$$\text{б)}\ y = \frac{1}{2} \ln \frac{1 + \operatorname{tg}\varphi \operatorname{tg} x}{1 - \operatorname{tg}\varphi \operatorname{tg} x}.$$

Например, при $R=1$ уравнение кривой а) принимает вид

$$y=\pm\ln(\cos x+\sqrt{1+\cos^2 x});$$

при $\varphi=\frac{\pi}{4}$ уравнение кривой б) принимает вид

$$y=\frac{1}{2}\ln\operatorname{ctg}\left(\frac{\pi}{4}-x\right).$$

На рис. 16 изображена криволинейная сетка на плоскости z, состоящая из кривых а) и б), переходящих при отображении на плоскость w соответственно в окружности с центром в начале и радиусом R и в лучи, выходящие из начала под углом φ к начальному лучу.

АЛФАВИТНЫЙ УКАЗАТЕЛЬ

Абель Н. Г. 448
Абеля лемма 448
Абсолютная величина 170
Амплитуда 102
Аналитическая функция 536, 543
— —, геометрический смысл 545
— —, представимость интегралом Коши 533
— —, представимость степенным рядом 539
— —, свойства 540
Аналитическое продолжение функции 543
Аргумент комплексного числа 503
Арифметическая прогрессия 143
— — порядка m 144
Арккосинус 130
Арксинус 128, 133
—, разложение в ряд 481
Арктангенс 131, 133
—, разложение в ряд 465
Ахиезер Н. И. 354

Бернулли И. 13
Бернштейн С. Н. 229, 302, 354
Бернштейна многочлены 229
— теорема 230
Бином Ньютона 472
Биномиальный ряд 473, 475
Больцано Б. 217, 312
Больцано-Вейерштрасса теорема 156, 159, 173
Больцано теорема 215, 245
Буняковский В. Я. 301

Валлиса формула 399
Вейерштрасс К. 228, 312
Вейерштрасса-Бернштейна теорема 229, 232
Вейерштрасса теоремы 195, 196, 218, 220, 269, 539
Величина бесконечно малая более высокого порядка 327
Величина переменная 299
Возрастание функции 35
Вторая производная 333
Второй дифференциал 336

Гармонический ряд 428
Гармоническое колебание простое 102
— — сложное 104
Гейне определение непрерывности функции 189
Геометрическая прогрессия 144, 428, 429
Гипербола 55
Гиперболические функции 86, 263, 507
Гиперболы высших порядков 56
— сопряжённые 73
Гиперсфера 265
Главные значения обратных тригонометрических функций 129, 130
Гомотетия 547
Гончаров В. Л. 354
Граница множества верхняя 287
— — нижняя 291
— последовательности верхняя 292
— — нижняя 292
График сложной функции, построение по точкам 24
— уравнения 17
— функции 17, 28
— —, вычислительный приём построения 18
— —, геометрический приём построения 19, 23
— —, построение по точкам 18, 22
Гюйгенса теорема 358

Даламбера признак сходимости ряда 435
Действительная часть комплексного числа 493
Декарт 299
Десятичные логарифмы 316
— —, точность вычисления 462
Дирихле 214

Дискриминант трёхчлена 47
Дифференциал 327, 330, 332
— второго порядка 336
— высшего порядка 336
—, геометрический смысл 329
—, применение к приближённым вычислениям 331
—, свойства 331
Дифференциальное исчисление, применение к исследованию функций 354
Дифференцирование 308, 330, 515
— обратной функции 318, 515
— сложной функции 322, 515
— функции нескольких переменных 337
Дифференцируемость функции 309, 514
Длина дуги кривой 417
— окружности 418
— промежутка 154
Дуга окружности 478

Евклидова метрика 265
Егоров Д. Ф. 301
Ермаков В. П. 437

Задача о касательной 304
— о массе стержня 377
— о плотности 306
— о площади криволинейной трапеции 379
— о пройденном пути 378
— о скорости 303
— приближения функции 228
Задачи на нахождение наибольшего значения функции 364
Замыкание множества 277
Золотарев Е. И. 354

Извлечение корней 476
Изолированный элемент множества 289
Инверсия 262
Интеграл, см. соответствующее название
Интегральная сумма 380, 382
Интегральное исчисленне, вычисление длины дуги 417
— —, — объёмов 411
— —, — площадей 409
— —, — площади поверхности вращения 419
Интегрирование 366, 424
— биномиальных дифференциалов 377
— методом подстановки 369, 371, 397
Интегрирование по частям 371, 397
— элементарных функций 373
Интегрируемость функции 520, 526
Интерполяционная задача 228
— формула Лагранжа 143

Кавальери принцип 410, 412
Касательная к кривой 305
Колмогоров А. Н. 301, 354
Комплексное число 493, 495
— —, показательная форма 506
— —, тригонометрическая форма 506
Континуум 279
Конформность отображений 547
Корень целой положительной степени 510
Коркин А. Н. 354
Косинус 96, 351, 482
— в комплексной области 501
—, разложение в ряд 471
Косинусоида 101
Коши О. 165
— интеграл 532
— теорема 342
— формула 342
— определение непрерывности функции 189
Коэффициенты степенного ряда 447
Кратность нуля аналитической функции 541
Крейн М. Г. 354
Кривая линия 517
Криволинейная трапеция 379
Криволинейный интеграл 531
Круговые функции 86, 96
Кубическая парабола 44, 69

Лагранжа интерполяционная формула 143
— теорема 342
— формула 342
Лейбниц 299
Лейбница теорема 437, 438
Лемма Абеля 448
Линия уровня 263
Лобачевский Н. И. 12, 214, 301, 312
Логарифм 87, 194, 249, 316
— в комплексной области 508, 530
—, приближённое вычисление 408
—, разложение в ряд 457
—, составление таблиц 408, 458, 463
—, функциональное свойство 89
Лозинский С. М. 354
Лопиталя правило 345
Лузин Н. Н. 301
Ляпунов А. М. 301

Максимум множества 287
— функции 36, 360
— — в расширенном смысле 37
Марков А. А. 301, 354
Марков В. А. 354
Масштаб отображения 546
Меньшов Д. Е. 301
Метрика в точечном пространстве 264
— в функциональном пространстве 268
— евклидова 265, 278
Минимум множества 291
— функции 36, 360
— — в расширенном смысле 37
Минковского неравенство 267
Мнимая часть комплексного числа 493
Многочлен биквадратный 51
— второй степени 46
— наименее отклоняющийся от функции 354
— первой степени 45
— рациональный 105, 106, 191
— Тейлора 348, 353
— третьей степени 48
Многочлены Бернштейна 229
— высшей степени 52
— тригонометрические 105, 106, 191
— Чебышева второго рода 109, 138
— —, минимальное свойство 134, 138
— — первого рода 109, 135, 136, 138
Множества гомеоморфные 283
Множество всюду плотное 238
— замкнутое 277
— компактное 270, 274
— ограниченное сверху 287
— — снизу 291
— определения функции 214
— производное 276
— пустое 275
— связное 278
— числовое 287
Модуль комплексного числа 497, 503
— перехода 90
Морера теорема 538
Муавра формула 505

Наибольшее значение функции 364
Наименьшее значение функции 364
Натуральные логарифмы 316, 408
— —, точность вычисления 459, 462
Неопределённый интеграл 367
— —, интегрирование методом замены переменной 369, 371
— —, — по частям 372
Непрерывность сложной функции 190
— функции 188, 190, 221, 310, 517
— — равномерная 221
Неравенство Минковского 267
— треугольника 266
Никольский С. М. 354
Нуль аналитической функции 540
— функции 37
— — кратности p 541
Ньютон И. 299
Ньютона бином 472
Ньютона-Лейбница формула 394

Область 520, 536
— определения функции 214
— существования аналитической функции 543
— сходимости последовательности 203
Обобщённая формула конечных приращений 342
Обратная функция 33, 217, 245, 318
— —, проблема существования 245
Обратные гиперболические функции 92, 513
— тригонометрические функции 128, 481, 512
— — —, главные значения 129
Общий член последовательности 142
Объединение множеств 275
Объём конуса 412
— тела 411, 415
— — вращения 416
— цилиндрического отрезка 414
— шара 412
Окрестность точки 155, 173
— точек $+\infty$ и ∞ 173
— элемента 264, 267
Окружность 72
Определённый интеграл 318, 518, 527
— —, вычисление с помощью неопределённого 393
— —, геометрический смысл 384
— —, дифференцирование по верхнему пределу 391
— —, интегрирование методом замены переменной 396, 529
— —, интегрирование по частям 397
— — как функция верхнего предела 391
— —, основные свойства 385, 519
— —, правило вычисления 395
— —, приближённое вычисление 400
— —, теорема об оценке 390
Остаток ряда 430
Остаточный член формулы Тейлора 350, 471
Остроградский М. В. 301, 377
Отклонение многочлена от функции 354
— функции от функции 222

Отображение 544
— взаимно однозначное 255
— гомеоморфное 283
— конформное 547
— —, примеры 547—552
— многозначное 255
— непрерывное 279
— однозначное 255
— плоскости на плоскость 262
— пространственное 260, 261
Ошибка приближённого равенства 460

Парабола 43, 47, 68
Параллельное перенесение 25
Первообразная функция 366
Пересечение множеств 275
Перестановка членов ряда 442
Период функции 30
Петровский И. Г. 301
Плотность стержня истинная 306, 377
— — средняя 306, 377
Площадь криволинейной трапеции 379, 384, 410
— круга 409
— поверхности вращения 419
— — шара 420
— шарового пояса 420
Поверхность уровня 263
Показательная функция 78, 193, 239
— — в комплексной области 501, 503
— —, периодичность 503
— —, функциональное свойство 82, 241, 249, 501
Полином 41
Полный дифференциал 339, 345
Полюс 526
Порядок алгебраической кривой 42
— тригонометрического многочлена 105
Последовательность бесконечная 141, 149
— возрастающая 196
— конечная 140, 149
— монотонная 197
— невозрастающая 196
— неограниченная 173, 272
— неубывающая 196
— ограниченная 152, 269
— — сверху 152
— — снизу 152
— равномерно сходящаяся 223
— расходящаяся 166
— сходящаяся 166, 202, 224, 269
— факториалов 145
— Фибоначчи 145
— функциональная 142
Последовательность числовая 142, 215
— —, геометрическое представление 148
— —, общее определение 149
— —, способы задания 142
Поссе К. А. 354
Правило Лопиталя 345
Предел 163, 165
— в несобственном смысле 173, 272
— множества 292
— — верхний 290
— — нижний 292
—, основные свойства 169
—, основные теоремы 169, 172, 497
— последовательности 163, 268, 271, 496
— — верхний 293
— — нижний 293
— — функций 203, 272
— — (числовой) 163, 165, 174
— при монотонном изменении 195, 199
— функции 187, 514
— — верхний 294
— — в точке 187
— — на бесконечности 176, 179
— — левый 185, 295
— — нижний 294
— — односторонний 185, 186
— — правый 186, 294
Предельная точка 158
Предельный переход 168, 210, 214
— элемент множества 275
— — последовательности 158, 269
Приближение функций многочленами 525
Привалов И. И. 301
Принцип Кавальери 410, 412
Проблема Чебышева 354
Произведение множеств 275
Производная 303, 308, 331, 513
— второго порядка 333
— высшего порядка 334, 337
—, геометрический смысл 309
— левосторонняя 311
— обратной функции 318
— правосторонняя 310
—, применение к выводу бинома Ньютона 334
—, — к доказательству неравенств 343, 357
Производные обратных тригонометрических функций 319
— элементарных функций, формулы и правила 312, 319, 324, 515
Произвольная постоянная 367
Промежуток 154, 215, 278, 279
— сходимости ряда 450
Протяжённость 259

Равенство комплексных чисел 459
Равномерная непрерывность функции 221
— сходимость последовательности 223
Радиус сходимости ряда 450
Разрыв функции в точке 181, 190
Ранг десятичной дроби 154
Расстояние 264
— между двумя точками на сфере 265
— — функциями 223, 268
Растяжение 25
Рациональная функция 41, 191, 494
Рациональный многочлен 41
Ремез Е. Я. 354
Риман 214
Ролля теорема 341
Ряд 425
— абсолютно сходящийся 441, 499
— в комплексной области 499
— гармонический 428
— знакочередующийся 437
— мажорантный к ряду 433
— положительный 431, 442
— —, свойство 437
— —, условие сходимости 432
— расходящийся 426
—, свойства 429, 430
— степенной 447
— строго положительный 431
— — —, признаки сходимости 434, 435, 436
— сходящийся 426, 499
— Тейлора 470

Сдвиг, трансляция 25
Сжатие 25
Симметрия осевая 25
— центральная 25
Синус 96, 116, 351, 482
— в комплексной области 501, 502
—, разложение в ряд 471
Синусоида 101
Синусоидальная кривая 102
Скорость изменения функции 309
Скорость точки истинная 303
— — средняя 303
Спрямление дуги окружности, правило Чебышева 478
Спрямляемая кривая 417
Степенная функция 42, 54, 69, 71, 93, 195, 249
Степенной ряд 447
— —, почленное дифференцирование 455
— —, — интегрирование 454
— —, свойства суммы 452, 454
Степень дробной рациональной функции 42
— многочлена 41, 67
Стрелка дуги 470
Сумма множеств 275
— ряда 426
Сфера 265
Сходимость абсолютная ряда 441, 443, 499
— знакочередующегося ряда, теорема Лейбница 437
— положительного ряда 442
— последовательности равномерная 223, 226
— — функций 202, 224
— ряда 426, 499
— —, необходимое условие 431
— —, признак Даламбера 435
— —, — сравнения рядов 433
— степенного ряда 448

Тангенс 111, 506
— половинного угла 115
Тейлора многочлен 348, 353
Тейлора ряд 470
— формула 348, 350
Теорема Бернштейна 230
— Больцано 215, 216, 245
— Больцано-Вейерштрасса 156, 159, 173
— Вейерштрасса 195, 196, 218, 539
— — о наибольшем значении 218, 220
— Вейерштрасса-Бернштейна 229, 232
— Гюйгенса 358
— Коши 342
— —, геометрический смысл 343
— Лагранжа 342
— —, геометрический смысл 343
— Лейбница 437, 438
— Морера 538
— об интегрируемости непрерывных функций 381
— о дифференцировании определенного интеграла по верхнему пределу 391
— о почленном дифференцировании степенного ряда 455, 456, 457
— — — интегрировании степенного ряда 454
— о равномерной непрерывности 220, 222, 281
— о среднем значении 387, 388
— Ролля 341, 342
— сложения гиперболических функций 86
— — показательных функций 82, 241, 248, 501

Теорема сложения произвольной функции 247
— — тригонометрических функций 101, 252, 482, 484, 505
— умножения произвольной функции 247
— Ферма 340, 360
— —, геометрический смысл 340
— Чебышева 377, 480
Тождество Эйлера 502, 504
Топологическое пространство 272
Точка внутренняя промежутка 154
— касания 305
— концевая 154
— перегиба 361
— стационарная 360
Точки экстремума функции 360
Тригонометрические функции 95, 111, 132, 192, 351, 501
— —, аналитическая теория 481, 484, 487
— —, выражение через показательную 504
Тригонометрический многочлен 105, 192

Убывание функции 35
Угловой коэффициент прямой 46
Умножение рядов 444, 499
Уравнение движения 304

Фаза гармонического колебания 102
Ферма П. 339
— теорема 340, 360
Фибоначчи последовательность 145
Формула Валлиса 399
— Коши 342
— Лагранжа 342, 343
— —, применения 343
— Муавра 505
— Ньютона-Лейбница 394
— средних прямоугольников 401
— Тейлора для многочлена 348
— — с остаточным членом 350, 352
Функции гиперболические, круговые, тригонометрические, обратные гиперболические и тригонометрические, элементарные, см. соответствующие названия
Функциональная зависимость 14
Функциональное пространство 267
— соответствие 255
Функциональные уравнения 247, 252
Функция 12, 211, 254, 299, 346, 496, 523, 544
Функция аналитическая, см. аналитическая функция
— алгебраическая 66
— — иррациональная 67
— —, примеры исследования 68
— возрастающая 35, 355
— голоморфная 537
— дифференцируемая 309
— дробная второй степени 58
— — линейная 56
— — рациональная 41, 64, 191
— e^x, разложение в ряд 352, 471
— интегрируемая 381, 520, 526
—, исследование на экстремум 361, 362
— класса K 360
— линейная 45, 46
— логарифмическая, см. логарифм
— многозначная 15, 211
— монотонная 37, 355
— невозрастающая 36
— непрерывная 188, 221, 517
— —, продолжение за пределы всюду плотного множества 237
— —, свойства 215
— неубывающая 36
— нечётная 29
— обратная 33, 217, 245, 318
— однозначная 15, 211
— первообразная 526
— периодическая 30
— показательная (экспоненциальная), см. показательная функция
— правильная 537
—, представимость интегралом Коши 533
—, приближение многочленами 525
—, признак монотонности 355
—, — постоянства 355
—, признаки существования экстремума 361
— промежутка 420, 424
— — аддитивная 421
— прямая 33
—, разложение в ряд 470, 535
— рациональная 41, 191, 494
— регулярная 535, 538
— $\operatorname{sgn} x$ 208
— сложная 322
— степенная, см. степенная функция
— точки 421, 424
— трансцендентная 78
— убывающая 35, 355
— целая рациональная 41
— чётная 29
Фурье 214

Целая часть x 211
Центр промежутка 154
— симметрии гиперболы 57

Частная производная 338
— —, геометрический смысл 338
Частичные суммы 426
Частота колебания 102
Чебышев П. Л. 138, 301, 353, 375
Чебышева многочлены 107, 109, 135
— правило приближённого спрямления дуги окружности 478
— проблема 354
— теоремы 375, 480
Число e 198, 352
Число π 399, 406, 466
Числовое множество 159, 287
Члены ряда 425

Эйлер Л. 13, 78, 198, 301
Эйлера тождество 502, 504
Экспоненциальная функция, см. показательная функция
Экстремум функции 36, 360
— —, правило нахождения 361
Элемент множества наибольший 287
— — наименьший 291
Элементарные функции 12, 195, 253, 374, 496, 514
Элементы множеств соответственные (гомологичные) 283
Эллипс, каноническое уравнение 72